DOCUMENT RESUME

ED 084 318 UD 013 879

AUTHOR	Kamin, Leon; Green, Winifred
TITLE	Transcripts From the Southern Regional Council's Symposium on Human Intelligence, Social Science, and Social Policy. Text of Dr. Kamin's Presentation Denying That Proof Exists That IQ Test Scores are Hereditary; How Testing Harms Children; IQ Tests as Instruments of Oppression--From Immigration Quotas to Welfare.
SPONS AGENCY	Southern Regional Council, Atlanta, Ga.
PUB DATE	73
NOTE	16p.; A reprint from "South Today"
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Ability Grouping; Educational Diagnosis; Environmental Influences; Genetics; *Heredity; *Intelligence Differences; *Intelligence Tests; Racial Differences; *Racial Discrimination; Research Methodology; Research Problems; Social Differences; Social Discrimination; Social Sciences; Student Grouping

ABSTRACT

This document comprises three presentations made on March 23 at a symposium sponsored by the Southern Regional Council focusing on Human Intelligence, Social Science and Social Policy. The first of the three parts of the document is the text of the principal presentation, made by Dr. Leon Kamin, Chariman of the Department of Psychology at Princeton University. The presentation first summarizes findings of his extensive research into the original studies on which some American social scientists have based writings which at least question whether environment has any effect on IQ test scores--suggesting that heredity may be the determinant. The second part of the document is the text of a presentation by Ms. Winifred Green, director of the Atlanta-based Southeastern Public Education Program of the American Friends Service Committee. The presentation argues that grouping reinforces the effects of years of discriminatory treatment in the education of black children--locking them into classroom situations where curriculum, materials, teacher expectation and the resulting stigmas and hopelessness are the same as, or some predict worse than, the days of separate but unequal schools. The third part is the text of Dr. Kamin's afternoon presentation on the history of the use in the United States to support repressive public policy. (Author/JM)

ED 084318

Transcripts From The Southern Regional Council's Symposium On Human Intelligence, Social Science, And Social Policy

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

A Reprint From

UD 013879

SOUTH TODAY
A Digest of Southern Affairs
Price: Twenty-five Cents

May–June, and July, 1973 Issues Vol. 4, No. 8 and 9

FILMED FROM BEST AVAILABLE COPY

Transcripts From
he Southern Regional Council's
Symposium
On Human Intelligence,
Social Science,
And Social Policy

TH.

REPRO
D FROM
ORIGIN
PINIONS
REPRE
TUTE OF
Y

A Reprint From

SOUTH TODAY
A Digest of Southern Affairs
Price: Twenty-five Cents

sues Vol. 4, No. 8 and 9 Atlanta, Georgia

FILMED FROM BEST AVAILABLE COPY

Text of Dr. Kamin's Presentation Denying That Proof Exists That IØ Test Scores Are Hereditar

On March 23, the Southern Regional Council sponsored a symposium with far-reaching implications on Human Intelligence, Social Science and Social Policy. The principal presentation was made by Dr. Leon Kamin, Chairman of the Department of Psychology at Princeton University.

Dr. Kamin first presented findings of his extensive research into the original studies on which some American social scientists have based writings which at least question whether environment has any effect on IQ test scores—suggesting that heredity may be the determinant. Dr. Kamin presented his findings in a quiet, almost off-hand manner. But what he said was startling and highly significant—in effect that there is no valid evidence at all to support the heredity assumption.

Since controversy over the heredity-environment question is one of the most important intellectual debates of our times, fraught with social policy implications and influence, we consider the following transcript of his presentation among the most important materials SOUTH TODAY has ever published. (Dr. Kamin spoke from notes; the transcript was tape-recorded. The transcript of his second presentation, on the history of the use of testing in America to support repressive public policy, will appear in a subsequent issue of SOUTH TODAY.)

My current impression—though I remain open to persuasion by the data—is that, upon detailed analysis, there is no evidence sufficiently strong to convince a reasonably prudent person to abandon the null hypothesis that there is no hereditary determination whatsoever of intelligence test scores. In arriving at that conclusion, I think it is important for you to know that it is not based on secondary sources or review articles, even though they appear in such distinguished journals as the *Atlantic Monthly*. If one goes back and looks at the actual experimental data which psychologists have accumulated on heritability of human intelligence, the evidence seems to crumble under one's probing finger so to speak.

Weight of Twins

Let me begin with an example of a particular study. Not because it is a terribly important one, but because it illustrates, to my mind at least, the way in which imperfect data are amplified in public journals which are currently made available to policy makers. If one reads Prof. Richard Herrnstein's celebrated article in the *Atlantic Monthly*, what he does there is first to review the evidence which claims to show that about 80 percent of the variation in intelligence is attributable to heredity. That leaves about 20 percent of the variance still to be accounted for.

And how do we account for that 20 percent which is not genetic? Well, said Prof. Herrnstein and I quote him: "The usual assumption that education and culture are critical is running into evidence that the physical environment, for example, early diet, might be more important. In fact, the twin studies that [Prof. Arthur] Jenson [educational psychologist at the University of California at Berkeley] surveyed showed that the single most important influence on IQ was not education or social environment, but something prenatal, as shown by the fact that the twin heavier at birth usually grew up with a higher IQ."

That's a rather astonishing finding. And it has clear policy implications. What the statement says is that insofar as there is an environmental effect at all on human intelligence, it's only the intrauterine environment during pregnancy which has any major effect. Therefore, perhaps we ought to feed welfare mothers with a special protein supplement during pregnancy. That might help the IQ's of their children, but there's no sense in wasting money on the cultural, social, or educational environment of these children. By the way, there's also no sense in feeding all of their fathers, if we are going to take this seriously, at least not to improve the kids' IQ's.

But what's the basis for this statement? Well, what Prof. Herrnstein has done in his article
earlier ar
Jensen, wr
lowing: "A
genetic va
and other
than diffe
logical en
favorablen
ment are
weight bet
in birth we
IQ differe

The sa
does Prof
cites a sin
he present
Scarr, who
of a numb
sented inf
twins wei
that the
invariably
The basic
must have
in the wor
blood sup

The fi
Jensen's
taken fro
break twi
For some
that each
For other
more tha
weighed l

xt of Dr. Kamin's Presentation Denying That Proof Exists t IQ Test Scores Are Hereditary

On March 23, the Southern Regional Council sponsored a symposium with far-reaching implications on Human Intelligence, Social Science and Social Policy. The principal presentation was made by Dr. Leon Kamin, Chairman of the Department of Psychology at Princeton University.

Dr. Kamin first presented findings of his extensive research into the original studies on which some American social scientists have based writings which at least question whether environment has any effect on IQ test scores—suggesting that heredity may be the determinant. Dr. Kamin presented his findings in a quiet, almost off-hand manner. But what he said was startling and highly significant—in effect that there is no valid evidence at all to support the heredity assumption.

Since controversy over the heredity-environment question is one of the most important intellectual debates of our times, fraught with social policy implications and influence, we consider the following transcript of his presentation among the most important materials SOUTH TODAY has ever published. (Dr. Kamin spoke from notes; the transcript was tape-recorded. The transcript of his second presentation, on the history of the use of testing in America to support repressive public policy, will appear in a subsequent issue of SOUTH TODAY.)

---

percent which is not genetic? Well, said Prof. Herrnstein and I quote him: "The usual assumption that education and culture are critical is running into evidence that the physical environment, for example, early diet, might be more important. In fact, the twin studies that [Prof. Arthur] Jenson [educational psychologist at the University of California at Berkeley] surveyed showed that the single most important influence on IQ was not education or social environment, but something prenatal, as shown by the fact that the twin heavier at birth usually grew up with a higher IQ."

That's a rather astonishing finding. And it has clear policy implications. What the statement says is that insofar as there is an environmental effect at all on human intelligence, it's only the intrauterine environment during pregnancy which has any major effect. Therefore, perhaps we ought to feed welfare mothers with a special protein supplement during pregnancy. That might help the IQ's of their children, but there's no sense in wasting money on the cultural, social, or educational environment of these children. By the way, there's also no sense in feeding all of their fathers, if we are going to take this seriously, at least not to improve the kids' IQ's.

But what's the basis for this statement? Well, what Prof. Herrnstein has done in his article is basically to paraphrase an earlier article by Prof. Jensen. Prof. Jensen, writing in 1970, had said the following: "A major proportion of the non-genetic variance is attributable to prenatal and other biological differences, rather than differences in the social or psychological environments. Differences in the favorableness of the intrauterine environment are reflected in differences in birth weight between twins. And the differences in birth weight are known to be related to IQ differences in twins."

The same thing—but on what basis does Prof. Jensen make this claim? He cites a single article by Sandra Scarr and he presents in his paper some data from Scarr, who had measured the intelligence of a number of identical twins and presented information about how much those twins weighed at birth. Scarr concludes that the heavier twin turns out almost invariably to grow up with a higher IQ. The basic idea being, the heavier twin must have had an easier and happier time in the womb, gotten more of the maternal blood supply, etc.

The first table is taken from Prof. Jensen's article, where he quotes data taken from Scarr. What Scarr did was to break twins up into three basic groups. For some twin pairs, it was true at birth that each weighed more than 2,500 grams. For other twin pairs, one twin weighed more than 2,500, but the other twin weighed less than 2,500 grams. And for

To page 2

Weight of Twins Studies Do Not Pr
As Claimed, Lack of Influence
Environment on a Child's IØ Sc

From page 1

the third type of twin, each twin weighed less than 2,500. Now the only data that Jensen presents are those which you see in Table I.

Table I

Mean IQ Difference in Favor of Heavier MZ Twins

Both Twins Greater Than 2500 grams	One Twin Less Than 2500 grams	Both Twins Less Than 2500 grams
4.9	13.3	6.4

And what that data show is that in the case where one twin weighed more than 2,500 grams and the other less, the heavier twin on the average has an IQ 13.3 points higher than the lighter twin. But notice that if both are reasonably fat and healthy, the heavier twin still has a higher IQ, but not that much higher, only 4.9 points; and similarily, if both twins are light in weight, the difference between them in IQ is not so large.

What Scarr concluded in her paper was, "the greater discrepancy in the group with one twin below 2,500 grams resulted from large differences in birth weight. Six of the eight pairs had birth differences of more than 500 grams. This large discrepancy may be caused by transfusion syndrome. One monozygotic twin bleeds into the other. The size of the IQ differences between co-twins was not affected by the absolute weight of the twins, but by the size of the weight differences between them."

Now she talks about single births, and concludes from these data: "One might conclude with justice, the bigger the better." The fatter the baby, presumably, the fatter the head, the fatter the brain, the higher the IQ. Strong biological reasoning.

But there are some problems with Scarr's data that I would like to call to your attention. First, it makes very good sense that she chose to work with twins, for the following reason. It is well known that there are social class and race differences in birth weight, in the United States at least. The average white child weighs more at birth than the average black child. The average upper-class child weighs more than the average lower-class child. So the advantage of working with identical twins is that when you are only 27 pairs, the average difference is only 6.3 IQ points. And that is a statistically significant difference.

In short, if one had reasonable caution, one would not have put together in a single analysis these two sets of twins, tested with two different so-called intelligence tests. That's a relatively small point. What about those three numbers in Table I that Prof. Jensen cites—4.9, 13.3, 6. something-or-other? If one does a traditional, simple, standard analysis for whether they differ significantly in a statistical sense, the answer is an unqualified no. There is no reason to conclude that the differences between those three numbers are due to anything other then chance. The differences between them are not, to use a technical term, statistically significant. That's a relatively minor point.

Now let's recall her logic, that if one twin is more than 2,500 grams and the other is less than 2,500 grams, that means there is a big birth weight difference between them. That means there's been something called transfusion syndrome. So, that in turn means that when there is a big birth weight difference, there will be a big IQ difference.

It is interesting to look at the eight twin pairs of whom it is true that one weighed more than 2,500 grams and the other less. As it happens, if you look at her data, for six of those pairs the difference in birth weight was in fact more than 500 grams. So that's "transfusion syndrone." And those twins differ on the average by 11 IQ points. But for the other two of her pairs, in that group, the difference in weight was much less than 500 grams. But with that small difference in birth weight, where presumably there has been no "bleeding" of one twin into the other in the womb, the difference in IQ turns out to be 20 points, not 11!

As you will see from Table I, no matter what category of twin you look at, the difference in IQ actually turns out to be larger for the twins who are close together in birth weight than for twins suffering from this hypothetical and never-observed bleeding and transfusion syndrome.

Scarr's logic says that the greater the difference in birth weight between a pair

Other Contrary Findings

But I don't think we ought to take it all that seriously, for the following reasons. Prof. Jensen brings Scarr's data into the very end of an article, during which he has been reviewing the evidence on separated identical twins. There are four such studies which he reviewed in that article. The only evidence that he gives his readers about birth weight in twins comes from this extra article by Scarr which he imports at the very end of his review article. That is peculiar. Because, as it happens, if one looks at the four studies which Prof. Jensen was reviewing, two of them had information on birth weight in twins. One of those studies was in England, by Shields, and Shields, for 61 pairs of twins, a larger sample than Scarr's, had asked the question — does the heavier twin at birth have the higher IQ? He concluded, "Birth weight is not associated with test score differences." Another of the four papers reviewed by Prof. Jensen was by Juel-Nielsen. She also raised the question whether there was any association between birth weight and IQ, and presented evidence to show no.

Those two studies that said no, Prof. Jensen did not choose to tell his readers about. The one that said yes, Prof. Jensen chose to cite. Now it's on the basis of this kind of empirical evidence that the readers of *Atlantic Monthly* are assured that psychological science has demonstrated that the major environmental variable affecting IQ is something prenatal. I think if we want to make policy on the basis of this kind of science, we had better not use science in making policy. I'm sorry to have taken so much time on that single example.

Identical Twin Studies

There is no possibility of reviewing this morning all the different types of evidence that psychologists have tried to bring together about the possible inheritance of intelligence. I would like to restrict my remarks to what by consensus is the strongest form of evidence, and also conceptually the simplest.

Primarily, I want to talk about the celebrated studies on separated identical

nt of Twins Studies Do Not Prove, Claimed, Lack of Influence of ronment on a Child's IQ Score

27 pairs, the average difference is only 6.3 IQ points. And that is a statistically significant difference.

In short, if one had reasonable caution, one would not have put together in a single analysis these two sets of twins, tested with two different so-called intelligence tests. That's a relatively small point. What about those three numbers in Table I that Prof. Jensen cites—4.9, 13.3, 6. something-or-other? If one does a traditional, simple, standard analysis for whether they differ significantly in a statistical sense, the answer is an unqualified no. There is no reason to conclude that the differences between those three numbers are due to anything other then chance. The differences between them are not, to use a technical term, statistically significant. That's a relatively minor point.

Now let's recall her logic, that if one twin is more than 2,500 grams and the other is less than 2,500 grams, that means there is a big birth weight difference between them. That means there's been something called transfusion syndrome. So, that in turn means that when there is a big birth weight difference, there will be a big IQ difference.

It is interesting to look at the eight twin pairs of whom it is true that one weighed more than 2,500 grams and the other less. As it happens, if you look at her data, for six of those pairs the difference in birth weight was in fact more than 500 grams. So that's "transfusion syndrone." And those twins differ on the average by 11 IQ points. But for the other two of her pairs, in that group, the difference in weight was much less than 500 grams. But with that small difference in birth weight, where presumably there has been no "bleeding" of one twin into the other in the womb, the difference in IQ turns out to be 20 points, not 11!

As you will see from Table I, no matter what category of twin you look at, the difference in IQ actually turns out to be larger for the twins who are close together in birth weight than for twins suffering from this hypothetical and never-observed bleeding and transfusion syndrome.

Scarr's logic says that the greater the difference in birth weight between a pair

Other Contrary Findings

But I don't think we ought to take it all that seriously, for the following reasons. Prof. Jensen brings Scarr's data into the very end of an article, during which he has been reviewing the evidence on separated identical twins. There are four such studies which he reviewed in that article. The only evidence that he gives his readers about birth weight in twins comes from this extra article by Scarr which he imports at the very end of his review article. That is peculiar. Because, as it happens, if one looks at the four studies which Prof. Jensen was reviewing, two of them had information on birth weight in twins. One of those studies was in England, by Shields, and Shields, for 61 pairs of twins, a larger sample than Scarr's, had asked the question — does the heavier twin at birth have the higher IQ? He concluded, "Birth weight is not associated with test score differences." Another of the four papers reviewed by Prof. Jensen was by Juel-Nielsen. She also raised the question whether there was any association between birth weight and IQ, and presented evidence to show no.

Those two studies that said no, Prof. Jensen did not choose to tell his readers about. The one that said yes, Prof. Jensen chose to cite. Now it's on the basis of this kind of empirical evidence that the readers of *Atlantic Monthly* are assured that psychological science has demonstrated that the major environmental variable affecting IQ is something prenatal. I think if we want to make policy on the basis of this kind of science, we had better not use science in making policy. I'm sorry to have taken so much time on that single example.

Identical Twin Studies

There is no possibility of reviewing this morning all the different types of evidence that psychologists have tried to bring together about the possible inheritance of intelligence. I would like to restrict my remarks to what by consensus is the strongest form of evidence, and also conceptually the simplest.

Primarily, I want to talk about the celebrated studies on separated identical

another a great deal, if they have been brought up in entirely different environments, then that can be seen reasonably as a consequence of the fact that they have the same heredity. Whereas, if they do not resemble one another in intelligence at all, it is clear that environment has largely determined their intelligence. The logic of this experiment is transparently simple to anybody.

As it happens, there are four studies in the psychological literature on separated twins. Everything I want to say now depends upon the correlation coefficient. Understand that if pairs of twins resemble one another perfectly, the correlation between them will be plus 1.00. If there was nothing but chance resemblance between twins, the correlation would be zero.

There are four studies of separated twins, and they have been graciously reviewed for us by no less an authority than Prof. Jensen in *Behavior Genetics* in 1970. The four studies were done by people named (1) Burt, (2) Shields, (3) Newman, Freeman and Holzinger, and (4) Juel-Nielsen. As you see in Table II they used rather different intelligence tests. The next column shows you the correlation coefficients for identical twins reared apart.

Table II

IQ Correlations For MZ Twins

Study	Test	Reared Apart
Burt	"Individual Test"	.86
Shields	Dominoes .77 + 2 x Mill Hill	
N.F. & H.	Stanford-Binet IQ	.67
Juel-Nielsen	Wechsler IQ	.62

But if you look at the third column, which is the critical one, you will see that there are some disagreements between the studies. The correlations that are reported for separated twins vary from .86 in the first case to .62 in Juel-Nielsen's. But nevertheless, it is clear that separated identical twins do indeed resemble one another in intelligence, when they are tested, to a very considerable degree.

Scrutiny of Burt's Data

The most important of these four studies is by Cyril Burt, for several rea-

conclude with justice, the bigger the better." The fatter the baby, presumably, the fatter the head, the fatter the brain, the higher the IQ. Strong biological reasoning.

But there are some problems with Scarr's data that I would like to call to your attention. First, it makes very good sense that she chose to work with twins, for the following reason. It is well known that there are social class and race differences in birth weight, in the United States at least. The average white child weighs more at birth than the average black child. The average upper-class child weighs more than the average lower-class child. So the advantage of working with identical twins is that when you are only comparing twins you have an absolute control for social class and for race. The pair of twins is obviously born in the same race and the same social class. So the basic logic makes sense.

So what Scarr had done was to take 25 pairs of twins she had studied herself, and combine those data with data from 27 other pairs of twins who had been studied by other people, namely Willerman and Churchill. And if one looks at the table of raw numbers in Scarr's paper, one discovers such facts as the following: The test which Scarr used to measure "IQ" was something called the Goodenough Draw-a-Person test. The test used by Willerman and Churchill was Wechsler Intelligence Scale for Children, the Performance Scale. Both had been combined in Scarr's study, which may sound fair enough, but there are some problems.

If you look, for example, at the average difference between a pair of twins, for Scarr's twenty-five pairs the average difference is 14.4 IQ points. For the other syndrome. And those twins differ on the average by 11 IQ points. But for the other two of her pairs, in that group, the difference in weight was much less than 500 grams. But with that small difference in birth weight, where presumably there has been no "bleeding" of one twin into the other in the womb, the difference in IQ turns out to be 20 points, not 11!

As you will see from Table 1, no matter what category of twin you look at, the difference in IQ actually turns out to be larger for the twins who are close together in birth weight than for twins suffering from this hypothetical and never-observed bleeding and transfusion syndrome.

Scarr's logic says that the greater the difference in birth weight between a pair of twins, the greater should be the difference in their IQ's. But when calculated, the correlation between the difference in birth weight and the difference in IQ turns out to be —.03. There is no correlation whatsoever between the difference in birth weight and the difference in IQ—none whatsoever. But I don't want to be unfair to Prof. Scarr. There is, in fact, one statistically significant effect in her data. That is the following: Suppose you don't ask, what is the difference in birth weight, but ask a very much simplier and blunter question. For any pair of twins, one is heavier and the other lighter. Forget how much heavier or how much lighter, and just ask the question, is it true that the heavier twin at birth—even if he's only one gram heavier—has the higher IQ, even if that IQ was 100 points higher? The answer is, yes. For her 52 pairs of twins, it is true that in 37 cases the heavier twin had a higher IQ than the lighter. For eight cases, the lighter twin had a higher IQ. In seven cases there were no differences. That *is* a statistically significant effect.

if we want to make policy on the basis of this kind of science, we had better not use science in making policy. I'm sorry to have taken so much time on that single example.

Identical Twin Studies

There is no possibility of reviewing this morning all the different types of evidence that psychologists have tried to bring together about the possible inheritance of intelligence. I would like to restrict my remarks to what by consensus is the strongest form of evidence, and also conceptually the simplest.

Primarily, I want to talk about the celebrated studies on separated identical twins. The point of these studies is really very simple. Identical twins, who are quite rare, are the only individuals in the world who have literally the same genes. As far as their heredity is concerned, they are identical. Not all twins, of course, are identical twins. But it can be determined with reasonable accuracy, by blood testing and various other procedures, whether twins are identical or not. If they are, then literally they have the same genes.

So in a way, nature provides the psychologist with a kind of experiment. Suppose you have a pair of identical twins, who for one reason or another are separated very early in childhood. Perhaps the mother dies, perhaps the family can't afford to bring up twins, whatever reason, one of these identical twins is reared in one home, in one environment, and the other twin is reared in a different home, a different environment.

If one can find such separated twins, and give them intelligence tests, then the interesting question to ask is, do they resemble one another in intelligence? To the extent that they do resemble one

syndrome. And those twins differ on the average by 11 IQ points. But for the other two of her pairs, in that group, the difference in weight was much less than 500 grams. But with that small difference in birth weight, where presumably there has been no "bleeding" of one twin into the other in the womb, the difference in IQ turns out to be 20 points, not 11!

As you will see from Table I, no matter what category of twin you look at, the difference in IQ actually turns out to be larger for the twins who are close together in birth weight than for twins suffering from this hypothetical and never-observed bleeding and transfusion syndrome.

Scarr's logic says that the greater the difference in birth weight between a pair of twins, the greater should be the difference in their IQ's. But when calculated, the correlation between the difference in birth weight and the difference in IQ turns out to be —.03. There is no correlation whatsoever between the difference in birth weight and the difference in IQ– none whatsoever. But I don't want to be unfair to Prof. Scarr. There is, in fact, one statistically significant effect in her data. That is the following: Suppose you don't ask, what is the difference in birth weight, but ask a very much simplier and blunter question. For any pair of twins, one is heavier and the other lighter. Forget how much heavier or how much lighter, and just ask the question, is it true that the heavier twin at birth—even if he's only one gram heavier—has the higher IQ, even if that IQ was 100 points higher? The answer is, yes. For her 52 pairs of twins, it is true that in 37 cases the heavier twin had a higher IQ than the lighter. For eight cases, the lighter twin had a higher IQ. In seven cases there were no differences. That *is* a statistically significant effect.

if we want to make policy on the basis of this kind of science, we had better not use science in making policy. I'm sorry to have taken so much time on that single example.

Identical Twin Studies

There is no possibility of reviewing this morning all the different types of evidence that psychologists have tried to bring together about the possible inheritance of intelligence. I would like to restrict my remarks to what by consensus is the strongest form of evidence, and also conceptually the simplest.

Primarily, I want to talk about the celebrated studies on separated identical twins. The point of these studies is really very simple. Identical twins, who are quite rare, are the only individuals in the world who have literally the same genes. As far as their heredity is concerned, they are identical. Not all twins, of course, are identical twins. But it can be determined with reasonable accuracy, by blood testing and various other procedures, whether twins are identical or not. If they are, then literally they have the same genes.

So in a way, nature provides the psychologist with a kind of experiment. Suppose you have a pair of identical twins, who for one reason or another are separated very early in childhood. Perhaps the mother dies, perhaps the family can't afford to bring up twins, whatever reason, one of these identical twins is reared in one home, in one environment, and the other twin is reared in a different home, a different environment.

If one can find such separated twins, and give them intelligence tests, then the interesting question to ask is, do they resemble one another in intelligence? To the extent that they do resemble one

Shields	Dominoes .77 + 2 x Mill Hill	
N.F. & H.	Stanford-Binet IQ	.67
Juel-Nielsen	Wechsler IQ	.62

But if you look at the third column, which is the critical one, you will see that there are some disagreements between the studies. The correlations that are reported for separated twins vary from .86 in the first case to .62 in Juel-Nielsen's. But nevertheless, it is clear that separated identical twins do indeed resemble one another in intelligence, when they are tested, to a very considerable degree.

Scrutiny of Burt's Data

The most important of these four studies is by Cyril Burt, for several reasons. First of all, he studied the largest number of pairs of separated twins — 53 pairs. Secondly, he reported the highest correlation of any study. And thirdly, as both Profs. Herrnstein and Jensen pointed out, Burt is the only investigator who presents quantitative data on the social or economic class of the homes in which the different twins were reared. That is important, for the following reason. Suppose it is the case that a pair of identical twins is placed in homes that resembled one another very closely. If the mother is a working-class mother, both twins might be placed in working class homes, separate homes but both working class homes. If the mother is a professional person and dies, and the twins go to two separate homes, both might be placed in professional homes. One must be certain that in fact the separated twins have been placed in different types of homes. It's important to have this information on what kinds of homes they were placed in. It's only in the Burt study that we were given

such information. So that makes the Burt study theoretically more important than any other.

Burt died a couple of years ago. He was English. He certainly was the major contributor by far to the data on heritability of human intelligence; a tremendous proportion of the data in this whole area, on how much relatives of different degrees of biological relatedness resemble each other in IQ, is due to the work of Burt and his students. He is the monumental figure in the area.

So I want to spend a little time looking at Prof. Burt. Let me say at once that whenever one reads Burt's papers, there

hygienic conditions." And what he means by cultural status is "educational and motivation background."

Now look at that .315. It now turns out that the correlation of .315 (or .32) is between *cultural* status and intelligence! Whereas we had been assured in 1943 that the correlation was between *economic* status and intelligence—definitely not cultural status. And we weren't told in 1943 that there had been any measurement at all of cultural status. There is a simple flat contradiction here, and this is not unrepresentative of Burt. I think the question has to be asked in all seriousness,

the intelligence correlation. And if you look at unrelated children reared together, there were more children again in 1966 than in 1955. Once again one discovers that most correlation coefficients haven't been changed appreciably by the additions of extra cases. The probability of this occurring is so astronomically small that one must conclude that there is no reasonable interpretation left—something is wrong. It just doesn't happen that way. You don't add 20 new pairs of twins and find all those correlations remaining identical.

Consider the case of non-identical twins. Burt reported on 172 pairs in

.944. By the
with 95 pairs
.944.

That's real
of intelligenc
ing such stat
Unless you
intelligence
can't appreci
data are.

The testing
tions quoted
gives us the
and child,
grandchild,
and so on. T

"...Whenever one reads Burt's papers there are certain ambiguities in th

are certain ambiguities in them. It's often very difficult to know precisely what he has done to whom and when. He presents correlation coefficients, but he doesn't give us very much information about what test he used, to whom he gave it, and so on.

I will give you some examples. In the first paper that he published which reported correlation coefficients for different types of relatives in 1943, Burt tells us: "Some of the inquiries have been published in London County Council reports or elsewhere. But the majority remain buried in typed memoranda or degree theses." It is going to be hard to get precise details about many of his studies, and often one wishes one had precise details for reasons such as the following.

If one starts in 1943, in that original paper, Burt reported a correlation between what he then called intelligence and "economic status." The higher the economic status, the higher the intelligence. The correlation was reported as .32. In that paper, Burt made a big point of stating that what he had measured (he doesn't tell us how he measured it) was economic status, *not* cultural status. He wants to make that point because, as he indicates, other investigators have shown a higher correlation than .32 between *cultural* status and intelligence. Therefore, to explain the fact that his correlation was only .32, he very specifically informs us that he measured *economic* status and not cultural status.

In 1956, with Burt referring to exactly the same study, we discover something. Now we get different sets of correlations. One set of correlation's is for something that Burt now calls the "crude test results," and the other is something that he calls "adjusted assessments." In the 1943

would you buy a used car from a salesman with this kind of record?

What Burt did was to go through life collecting cases. He got interested in this problem relatively early in his life. And for example, at one point he would write a paper in which he might have a hundred pairs of twins, and over the next five years he might discover 25 more pairs, and five years later he would write a new paper, now based on 125 pairs of twins. He simply kept on adding cases during a very lengthy career. He died in his eighties, and was still working away.

Consider the case of siblings reared apart. In 1955, Burt had reported on 131 pairs of siblings reared apart. By 1966, the number of pairs had increased to 151. There were some remarkable coincidences, as shown in Table III. The addi-

Table III

Correlations From Burt

	Siblings Reared Apart		DZ Twins Reared Together	
	1955	1966	1955	1966
Intelligence				
Group Test	.441	.412	.542	.552
Individual Test	.463	.423	.526	.527
Assessment	.517	.438	.551	.453
School Attainment				
Reading, Spelling	.490	.490	.915	.919
Arithmetic	.563	.563	.748	.748
General	.526	.526	.831	.831
Physical				
Height	.536	.536	.472	.472
Weight	.427	.427	.586	.586

tion of 20 new pairs of twins didn't change to the third decimal place the correlations either for weight or height or general school attainment or arithmetic or English or spelling. The addition of 20 new pairs of twins didn't affect one of those correlation coefficients even to the third

1955. By 1966, 45 of those pairs had disappeared. We now have only 127. But that's all right; it doesn't change the correlation coefficient for height and weight and general school attainment and arithmetic. It makes no difference, in Prof. Burt's data, whether you add pairs or lose pairs, you get the same thing all the time. That's rather surprising, and it doesn't inspire confidence.

Now, this has to do with the separated identical twin studies. Prof. Burt evidently had a marvelous test, a group test of intelligence, which produced some really remarkable data. Concern yourself with only the first column and the fourth column of Table IV. The first column

Table IV

Correlations For MZ Twins*

Separated			Reared Together	
Group Test of Intelligence			Group Test of Intelligence	
1955:	.771	N=21	.944	N=83
1958(A):	.771	N="over 30"	.944	N= ?
1958(B):	.778	N=42	.936	N= ?
1966:	.771	N=53	.944	N=95

*First and fourth columns referred to in text

gives you the correlation coefficients for separated identical twins. At different points in Burt's career he reported on progressively larger and larger numbers of such twins. In his 1955 paper, when he had only 21 pairs of twins, the correlation for this marvelous group test of intelligence was .771. By early 1958, he had found at least nine more pairs. Now he had over 30, but the correlation coefficient was still .771. And for 53 pairs in 1966, it is still .771!

Now this same marvelous group intelligence test was also given to non-separated identical twins and notice that with 83 pairs he had .944. Again, for an

those correla
genetic sens
herited. Burt
and it's very
to children,
school psych
room anywa
when you gi

But how
gence of a
1957 tell us
of our surv
ually attai
sample of p
of standardi
type." Reflec
says. Most
rental intell
were gotten
talk to the
sessment of
gence is. W
ment, you
child's IQ
what kin
reassures us
representati
intelligence

Burt and
correlations
child, and
and grandch
us in that pa
the intellige
procedures
have already
publications.
ious publica
reads that p
a footnote
information:
parents, we
interviews. B
cases, an ope
employed."

hygienic conditions." And what he means by cultural status is "educational and motivation background."

Now look at that .315. It now turns out that the correlation of .315 (or .32) is between *cultural* status and intelligence! Whereas we had been assured in 1943 that the correlation was between *economic* status and intelligence—definitely not cultural status. And we weren't told in 1943 that there had been any measurement at all of cultural status. There is a simple flat contradiction here, and this is not unrepresentative of Burt. I think the question has to be asked in all seriousness,

the intelligence correlation. And if you look at unrelated children reared together, there were more children again in 1966 than in 1955. Once again one discovers that most correlation coefficients haven't been changed appreciably by the additions of extra cases. The probability of this occurring is so astronomically small that one must conclude that there is no reasonable interpretation left—something is wrong. It just doesn't happen that way. You don't add 20 new pairs of twins and find all those correlations remaining identical.

Consider the case of non-identical twins. Burt reported on 172 pairs in

.944. By the time he concludes his career with 95 pairs of those twins, he's still at .944.

That's really a remarkable group test of intelligence that's capable of providing such stable data. Truly remarkable. Unless you have worked yourself with intelligence tests and correlations, you can't appreciate how remarkable Burt's data are.

The testing literature is full of correlations quoted from Burt, in which he gives us the correlation between parent and child, between grandparent and grandchild, between uncle and nephew, and so on. The point is to show that all

...never one reads Burt's papers, ... are certain ambiguities in them."

would you buy a used car from a salesman with this kind of record?

What Burt did was to go through life collecting cases. He got interested in this problem relatively early in his life. And for example, at one point he would write a paper in which he might have a hundred pairs of twins, and over the next five years he might discover 25 more pairs, and five years later he would write a new paper, now based on 125 pairs of twins. He simply kept on adding cases during a very lengthy career. He died in his eighties, and was still working away.

Consider the case of siblings reared apart. In 1955, Burt had reported on 131 pairs of siblings reared apart. By 1966, the number of pairs had increased to 151. There were some remarkable coincidences, as shown in Table III. The addi-

Table III

Correlations From Burt

	Siblings Reared Apart		DZ Twins Reared Together	
	1955	1966	1955	1966
Intelligence				
Group Test	.441	.412	.542	.552
Individual Test	.463	.423	.526	.527
Assessment	.517	.438	.551	.453
School Attainment				
Reading, Spelling	.490	.490	.915	.919
Arithmetic	.563	.563	.748	.748
General	.526	.526	.831	.831
Physical				
Height	.536	.536	.472	.472
Weight	.427	.427	.586	.586

tion of 20 new pairs of twins didn't change to the third decimal place the correlations either for weight or height or general school attainment or arithmetic or English or spelling. The addition of 20 new pairs of twins didn't affect one of those correlation coefficients even to the third

1955. By 1966, 45 of those pairs had disappeared. We now have only 127. But that's all right; it doesn't change the correlation coefficient for height and weight and general school attainment and arithmetic. It makes no difference, in Prof. Burt's data, whether you add pairs or lose pairs, you get the same thing all the time. That's rather surprising, and it doesn't inspire confidence.

Now, this has to do with the separated identical twin studies. Prof. Burt evidently had a marvelous test, a group test of intelligence, which produced some really remarkable data. Concern yourself with only the first column and the fourth column of Table IV. The first column

Table IV

Correlations For MZ Twins*

Separated			Reared Together	
Group Test of Intelligence			Group Test of Intelligence	
1955:	.771	N=21	.944	N=83
1958(A):	.771	N="over 30"	.944	N= ?
1958(B):	.778	N=42	.936	N= ?
1966:	.771	N=53	.944	N=95

*First and fourth columns referred to in text

gives you the correlation coefficients for separated identical twins. At different points in Burt's career he reported on progressively larger and larger numbers of such twins. In his 1955 paper, when he had only 21 pairs of twins, the correlation for this marvelous group test of intelligence was .771. By early 1958, he had found at least nine more pairs. Now he had over 30, but the correlation coefficient was still .771. And for 53 pairs in 1966, it is still .771!

Now this same marvelous group intelligence test was also given to non-separated identical twins and notice that with 83 pairs he had .944. Again, for an

those correlations make perfect biological genetic sense, that intelligence is inherited. Burt was a school psychologist, and it's very easy to give intelligence tests to children, particularly if you are a school psychologist. They are in a classroom anyway, they're a captive audience when you give them a test.

But how do you measure the intelligence of adults? Burt and Howard in 1957 tell us the following: "But in each of our surveys assessments were individually attained for a representative sample of parents, checked for purposes of standardization by tests of the usual type." Reflect for a moment on what that says. Most of the measurements for parental intelligence, or adult intelligence, were gotten by *assessment*. That is, you talk to the adult and you make an assessment of what you think his intelligence is. When you make that assessment, you probably know what his child's IQ is, and you certainly know what kind of home he lives in. Prof. Burt reassures us that he does give, at least to a representative sample of adults, a real intelligence test of the usual type.

Burt and Howard in 1966 reported correlations for 963 pairs of parent and child, and for 321 pairs of grandparent and grandchild. The only thing they tell us in that paper about how they measured the intelligence of the adults was, "The procedures employed and results obtained have already been described in previous publications." And they cite one previous publication. If one goes back and reads that previous publication, one finds a footnote which gives us this bit of information: "For the assessment of the parents, we relied chiefly on personal interviews. But in doubtful and borderline cases, an open or a camouflaged test was employed." Now, the picture of Prof.

of stating that what he had measured (he doesn't tell us how he measured it) was economic status, *not* cultural status. He wants to make that point because, as he indicates, other investigators have shown a higher correlation than .32 between *cultural* status and intelligence. Therefore, to explain the fact that his correlation was only .32, he very specifically informs us that he measured *economic* status and not cultural status.

In 1956, with Burt referring to exactly the same study, we discover something. Now we get different sets of correlations. One set of correlation's is for something that Burt now calls the "crude test results," and the other is something that he calls "adjusted assessments." In the 1943 paper we didn't know much about this distinction. What it means is this. You give someone an intelligence test and you score it, that's the crude test mark. But if you feel there is some inaccuracy in the test, it didn't give a perfect measure, then you adjust it. You take the test score and if it doesn't seem accurate to you, you adjust it.

It turns out, as you will see, that the correlation which in 1943 had been reported as between intelligence and economic status was really the correlation between the adjusted assessment (about which we hear for the first time in 1956) and "socio-economic status." But now in 1957, Burt has another paper where he refers once again to exactly the same study, and look at what has happened. It turns out that he evidently had all along not only a measure of economic status, but a measure of cultural status. Burt's use of words is getting loose now, and it turns out what he means by economic status is "material, i.e. financial and

	1955	1966	1955	1966
Intelligence				
Group Test	.441	.412	.542	.552
Individual Test	.463	.423	.526	.527
Assessment	.517	.438	.551	.453
School Attainment				
Reading, Spelling	.490	.490	.915	.919
Arithmetic	.563	.563	.748	.748
General	.526	.526	.831	.831
Physical				
Height	.536	.536	.472	.472
Weight	.427	.42	.586	.586

*First and fourth columns referred to in text

tion of 20 new pairs of twins didn't change to the third decimal place the correlations either for weight or height or general school attainment or arithmetic or English or spelling. The addition of 20 new pairs of twins didn't affect one of those correlation coefficients even to the third decimal place! It did slightly change gives you the correlation coefficients for separated identical twins. At different points in Burt's career he reported on progressively larger and larger numbers of such twins. In his 1955 paper, when he had only 21 pairs of twins, the correlation for this marvelous group test of intelligence was .771. By early 1958, he had found at least nine more pairs. Now he had over 30, but the correlation coefficient was still .771. And for 53 pairs in 1966, it is still .771!

Now this same marvelous group intelligence test was also given to non-separated identical twins and notice that with 83 pairs he had .944. Again, for an unspecified number of pairs in 1958,

George L. Walker III

Dr. Kamin Discusses Presentation With SRC Staffer Alexis Herman

	1955	1966	1955	1966
Intelligence				
Group Test	.441	.412	.542	.552
Individual Test	.463	.422	.526	.527
Assessment	.517	.438	.551	.453
School Attainment				
Reading, Spelling	.490	.490	.915	.919
Arithmetic	.563	.563	.748	.748
General	.526	.526	.831	831
Physical				
Height	.536	.536	.472	.472
Weight	.427	.427	.586	.586

*First and fourth columns referred to in text

tion of 20 new pairs of twins didn't change to the third decimal place the correlations either for weight or height or general school attainment or arithmetic or English or spelling. The addition of 20 new pairs of twins didn't affect one of those correlation coefficients even to the third decimal place! It did slightly change gives you the correlation coefficients for separated identical twins. At different points in Burt's career he reported on progressively larger and larger numbers of such twins. In his 1955 paper, when he had only 21 pairs of twins, the correlation for this marvelous group test of intelligence was .771. By early 1958, he had found at least nine more pairs. Now he had over 30, but the correlation coefficient was still .771. And for 53 pairs in 1966, it is still .771!

Now this same marvelous group intelligence test was also given to non-separated identical twins and notice that with 83 pairs he had .944. Again, for an unspecified number of pairs in 1958,

George L. Walker III

Dr. Kamin Discusses Presentation With SRC Staffer Alexis Herman

intelligence test of the usual type."

Burt and Howard in 1966 reported correlations for 963 pairs of parent and child, and for 321 pairs of grandparent and grandchild. The only thing they tell us in that paper about how they measured the intelligence of the adults was, "The procedures employed and results obtained have already been described in previous publications." And they cite one previous publication. If one goes back and reads that previous publication, one finds a footnote which gives us this bit of information: "For the assessment of the parents, we relied chiefly on personal interviews. But in doubtful and borderline cases, an open or a camouflaged test was employed." Now, the picture of Prof. Burt giving a camouflaged test of intelligence to a grandparent in London is perhaps an entertaining one, but it does not inspire scientific confidence in most minds.

And note also the magical transmutation. Remember that each of the following quotes refers to precisely the same study. In 1955, it is "doubtful and borderline cases," but by 1957 in Prof. Burt's memory, this has become "a representative sample of parents." In 1955 the tests were "open or camouflaged." But by 1957, in the mist of Prof. Burt's memories, they had become "tests of the usual type." This doesn't really inspire a great deal of confidence.

Nevertheless, all of our introductory textbooks in psychology, all our review articles in *Atlantic Monthly* and other learned journals will quote the correlations reported by Burt, giving the impression that psychological science has demonstrated that the correlations in intelligence test scores for relatives is this and that and that. The poverty of the

To page 4

Four Studies of Separated Identical Twin
Are Shot Through With Methodological I

From page 3

empirical and procedural background doesn't come through when you read reviews of what is alleged to be shown by these data.

How about this question of using assessments rather than using test scores? Let me read a couple of quotations, from Prof. Burt. Burt writes in 1958, "The data were secured in the course of surveys. The final assessments for the children were obtained by submitting the marks from the group tests to the judgment of the teachers. Where the teacher disagreed with the verdict of the marks, the child was interviewed personally and subjected to further tests, often on several successive occasions. The assessments for the adult members of the families were naturally far less accurate." That is, if a particular child gets an intelligence test score which doesn't seem to fit, he can be tested and has been tested by Prof. Burt a second time and a third time, and if necessary, a fourth time. How much credibility are we to attach to correlation coefficients based not upon the standardized administration of an intelligence test but upon a repeatedly adjusted assessment made by teachers or by Prof. Burt? The most detailed explanation that Burt makes of this assessment procedure is as follows: "We are perfectly willing to admit [this is Burt in 1956] that as a means of estimating genotypic differences, [that is, differences due to the genes] tests are highly fallible instruments, far less trustworthy than the judgments of teachers. An officially appointed psychologist possesses the authority to extend or repeat his tests and his interviews, and to require from teachers further assistance. The data have been secured in this way and having satisfied ourselves that by these means we can reduce the disturbing effects of environment to relatively slight proportions, we have gone on . . . ," etc.

So what Prof. Burt is doing is deliberately removing from his numbers "the disturbing effect of environment." Because he is interested only in the alleged effect of the genes, he removes the disturbing effect of environment on the test scores by adjusting the scores in terms of comments and criticisms made by the teachers. What's very peculiar about this is that these adjusted scores were then mailed by Prof. Burt to, among

Now, anybody is entitled to change his mind, but it is instructive to ask in what context Prof. Burt made that early statement. The earlier quotation was in 1943, when the debate was already on in England about who was to be allowed to attend university and how children were to be streamed in the public schools. At that point, Prof. Burt was arguing essentially the following: We should segregate early on in their careers the genetically bright. They are the ones who will be allowed to go on to universities. And to find out who is genetically bright, to find out who is going to be allowed to go into what classroom, he advocated use of the test score because it's a better estimator of genes. Don't allow the teacher to influence student placement!

When it suited Prof. Burt's purpose to keep the teacher out of socially important decisions, he argued that the test was a marvelous measure of hereditary intelligence and the teacher was not to be trusted. But later in his career, Prof. Burt had test data which he tried to fit to a theoretical model of inheritance, and the data didn't fit that well. Then he turned to the teacher to help him out. Now the teacher was the better estimator of the genotype.

Twin Study Inconsistencies

We finally get to the separated twins study itself. It would be reasonable to ask what intelligence test did Burt give to the separated twins? The answer is, I'm afraid, we do not know. Prof. Jensen does know. He informed us very confidently in 1970, talking about Burt's data, "their IQ's were obtained from an individual test, the English adaptation of the Stanford-Binet." They were not, by the way.

Burt in 1966 had said the following: "The tests employed have been fully described elsewhere. They consisted of (1) a group test of intelligence containing both non-verbal and verbal items; (2) an individual test, the London revision of the Terman-Binet scale, used primarily for standardization, and for doubtful cases, (3) a set of performance tests standardized by Miss Gaw in 1925." Of course the test results were submitted to teachers, and so on. Now the interesting thing

one of those seven, if any, he used. In fact he couldn't have used any of those seven, because he says clearly, "All the group tests in the present category are linguistic in form." But elsewhere he has informed us that his marvelous intelligence test had both verbal and non-verbal items.

Now, about those seven group tests which he gives us, he tells us such things as the following: "Complete tables of age norms would be unnecessary or even misleading. I give only rough averages, calculated regardless of sex. I have not thought it worth the necessary time and space to elaborate or print a set of standardized instructions as to procedures or marking."

Nevertheless, that's the test which repeatedly gives us that remarkably stable correlation of .771, no matter how large the number of twins. This is really a terribly important question, for the following reason: Intelligence tests in theory have been constructed in such a way that, for each sex the average IQ is 100, and for each age, the average IQ is 100. To the degree that the test might possibly not be perfect, we are going to find serious problems. Consider the following: Any pair of identical twins is necessarily of the same sex and of the same age. Suppose you are using a test where, for example, men score higher than women, or where older children score higher than younger children. That means that when you compare twins, they may resemble one another simply because they are the same sex or simply because they are the same age. And not because they have the same genes. Unless one is assured that the test which one is using in these studies does not give different scores for different ages, or for different sexes, the studies simply hold no water. I have looked into Miss Gaw's performance test, standardized in 1925. She informs us that there are "striking" differences between the sexes in the performance test that Prof. Burt used.

Recall that Burt's is the only study that gives us data on the social and economic status of the homes in which separated twins were placed. By 1958 Burt had reported on 42 pairs of separated twins. In 1958 we were told that at least four children of high-professional

ies of Separated Identical Twins Through With Methodological Flaws

Now, anybody is entitled to change his mind, but it is instructive to ask in what context Prof. Burt made that early statement. The earlier quotation was in 1943, when the debate was already on in England about who was to be allowed to attend university, and how children were to be streamed in the public schools. At that point, Prof. Burt was arguing essentially the following: We should segregate early on in their careers the genetically bright. They are the ones who will be allowed to go on to universities. And to find out who is genetically bright, to find out who is going to be allowed to go into what classroom, he advocated use of the test score because it's a better estimator of genes. Don't allow the teacher to influence student placement!

When it suited Prof. Burt's purpose to keep the teacher out of socially important decisions he argued that the test was a marvelous measure of hereditary intelligence and the teacher was not to be trusted. But later in his career, Prof. Burt had test data which he tried to fit to a theoretical model of inheritance, and the data didn't fit that well. Then he turned to the teacher to help him out. Now the teacher was the better estimator of the genotype.

Twin Study Inconsistencies

We finally get to the separated twins study itself. It would be reasonable to ask what intelligence test did Burt give to the separated twins? The answer is, I'm afraid, we do not know. Prof. Jensen does know. He informed us very confidently in 1970, talking about Burt's data, "their IQ's were obtained from an individual test, the English adaptation of the Stanford-Binet." They were not, by the way.

Burt in 1966 had said the following: "The tests employed have been fully described elsewhere. They consisted of (1) a group test of intelligence containing both non-verbal and verbal items; (2) an individual test, the London revision of the Terman-Binet scale, used primarily for standardization, and for doubtful cases, (3) a set of performance tests standardized by Miss Gaw in 1925." Of course the test results were submitted to teachers, and so on. Now the interesting thing one of those seven, if any, he used. In fact he couldn't have used any of those seven, because he says clearly, "All the group tests in the present category are linguistic in form." But elsewhere he has informed us that his marvelous intelligence test had both verbal and non-verbal items.

Now, about those seven group tests which he gives us, he tells us such things as the following: "Complete tables of age norms would be unnecessary or even misleading. I give only rough averages, calculated regardless of sex. I have not thought it worth the necessary time and space to elaborate or print a set of standardized instructions as to procedures or marking."

Nevertheless, that's the test which repeatedly gives us that remarkably stable correlation of .771, no matter how large the number of twins. This is really a terribly important question, for the following reason: Intelligence tests in theory have been constructed in such a way that, for each sex the average IQ is 100, and for each age, the average IQ is 100. To the degree that the test might possibly not be perfect, we are going to find serious problems. Consider the following: Any pair of identical twins is necessarily of the same sex and of the same age. Suppose you are using a test where, for example, men score higher than women, or where older children score higher than younger children. That means that when you compare twins, they may resemble one another simply because they are the same sex or simply because they are the same age. And not because they have the same genes. Unless one is assured that the test which one is using in these studies does not give different scores for different ages, or for different sexes, the studies simply hold no water. I have looked into Miss Gaw's performance test, standardized in 1925. She informs us that there are "striking" differences between the sexes in the performance test that Prof. Burt used.

Recall that Burt's is the only study that gives us data on the social and economic status of the homes in which separated twins were placed. By 1958 Burt had reported on 42 pairs of separated twins. In 1958 we were told that at least four children of high professional

of the relations between politics and psychology than some of his American contemporaries. He wrote. "In psychology, as in politics, the pendulum of the fashion swings to and fro. During the nineteenth century, the associationists preached an egalitarian doctrine, and three reform bills were passed." Burt's work is not likely to lead to any new reform bills.

Other Twin Studies

Burt's is only one of the four studies; what about the others? The second was also done in England, by Shields. It is the second largest study, with approximately 40 pairs of twins. Where Burt provides us with no procedural details, Shields provides us with all the possible information we could ask for. I'm going to skip over some of the more technical procedural details.

First of all, we can get from Prof. Shields an indication of what psychologists mean when they talk about separated identical twins. I'm quoting from his Appendix: These three cases were "separated twins." Benjamin and Ronald were separated at nine months. "Both brought up in the same fruit-growing village, Ben by the parents, Ron by the grandmother. They were at school together, they have continued to live in the same village." Ben and Ron were 52-years-old when they were brought from their village to London, to be tested by Shields, who then found out that they resembled one another.

Jessica and Winifred were separated at three months. "Brought up within a few hundred yards of one another. Attracted to each other at the age of two, but meetings were not encouraged. Told that they were twins, after the girls discovered it for themselves, gravitated to one another at school. They play together quite a lot at school and during the evenings. Jessica often goes to tea with Winifred. They were never apart and wanted to sit at the same desk." This is a pair of separated twins, eight-years-old when, after this terrific separation, Shields discovered that they resembled one another.

Burton and Christopher were separated at birth. "The paternal aunts decided to take one twin each and brought them up amiably living next door to one another

from teachers further assistance. The data have been secured in this way and having satisfied ourselves that by these means we can reduce the disturbing effects of environment to relatively slight proportions, we have gone on . . . ," etc.

So what Prof. Burt is doing is deliberately removing from his numbers "the disturbing effect of environment." Because he is interested only in the alleged effect of the genes, he removes the disturbing effect of environment on the test scores by adjusting the scores in terms of comments and criticisms made by the teachers. What's very peculiar about this is that these adjusted scores were then mailed by Prof. Burt to, among others, Prof. Jensen and Prof. Shockley in the United States, who had a great deal of fun with those numbers. Prof. Jensen has used those numbers — from which the disturbing effects of environmental influence had been systematically removed — to estimate the proportion of variance in intelligence attributable to heredity vs. environment. It is a kind of comedy, but the possible effects on social policy are not so funny.

Prof. Burt was not always so confident in the teacher as the best estimator of the genetic portion of intelligence. Burt says in 1956 that the test is far less trustworthy than the teacher. Thirteen years earlier he had written, "But in regard to innate general ability there can be no question. The unaided judgments even of the most experienced teachers, shrewd as they are in many cases, are nevertheless far less trustworthy in the long run than results obtained with properly administered intelligence tests." This is in absolute flat contradiction to what he then wrote 15 years later.

dently in 1970, talking about Burt's data, "their IQ's were obtained from an individual test, the English adaptation of the Stanford-Binet." They were not, by the way.

Burt in 1966 had said the following: "The tests employed have been fully described elsewhere. They consisted of (1) a group test of intelligence containing both non-verbal and verbal items; (2) an individual test, the London revision of the Terman-Binet scale, used primarily for standardization, and for doubtful cases, (3) a set of performance tests standardized by Miss Gaw in 1925." Of course the test results were submitted to teachers, and so on. Now the interesting thing is that in 1958 Prof. Burt had replied to an earlier critic and said the following about the "individual test" used on the twins: "The figures he quotes from my own research were based on a non-verbal test of performance type."

So in 1958, Prof. Burt blandly informs us that the individual test that he used with the twins was a non-verbal test of the performance type. In 1966, he blandly informs us that it was the London revision of the Terman-Binet, which is a *verbal* test *not* of the performance type. Another flat contradiction.

We simply cannot discover from Prof. Burt what tests were given by whom to whom and under what circumstances. He does, as you see, refer us to other sources to find out about this marvelous group test of intelligence which he's given. He gives two sources, and in one source there are no group intelligence tests at all. In the other source there are no fewer than seven group intelligence tests which he had presented. We can't figure out which

the same age. And not because they have the same genes. Unless one is assured that the test which one is using in these studies does not give different scores for different ages, or for different sexes, the studies simply hold no water. I have looked into Miss Gaw's performance test, standardized in 1925. She informs us that there are "striking" differences between the sexes in the performance test that Prof. Burt used.

Recall that Burt's is the only study that gives us data on the social and economic status of the homes in which separated twins were placed. By 1958 Burt had reported on 42 pairs of separated twins. In 1958 we were told that at least four children of high professional families—"class 1" in his study — had been reared in orphanages. But in 1966, reporting on 53 pairs, zero children of class 1 professionals had been reared in orphanages and only two of clerical workers, the next class down, had been reared in orphanages. How did four out of 42 become zero out of 53? Flat contradiction. Furthermore, Prof. Burt after 1966 mailed this marvelous data on social class to Prof. Jensen. Due to the kindness of Prof. Jensen I have a copy of the data which Prof. Burt mailed. Those data in turn contradict 1966 data, which in turn contradict the 1958 data.

If you look for mathematical impossibilities in Burt's work they assail you from every side. I really have to conclude, and I think I have presented enough evidence so that most reasonable people can follow me, that one simply cannot take Prof. Burt's numbers seriously.

I will say one thing for Prof. Burt: He was much more conscious and aware

dently in 1970, talking about Burt's data, "their IQ's were obtained from an individual test, the English adaptation of the Stanford-Binet." They were not, by the way.

Burt in 1966 had said the following: "The tests employed have been fully described elsewhere. They consisted of (1) a group test of intelligence containing both non-verbal and verbal items; (2) an individual test, the London revision of the Terman-Binet scale, used primarily for standardization, and for doubtful cases, (3) a set of performance tests standardized by Miss Gaw in 1925." Of course the test results were submitted to teachers, and so on. Now the interesting thing is that in 1958 Prof. Burt had replied to an earlier critic and said the following about the "individual test" used on the twins: "The figures he quotes from my own research were based on a non-verbal test of performance type."

So in 1958, Prof. Burt blandly informs us that the individual test that he used with the twins was a non-verbal test of the performance type. In 1966, he blandly informs us that it was the London revision of the Terman-Binet, which is a *verbal* test *not* of the performance type. Another flat contradiction.

We simply cannot discover from Prof. Burt what tests were given by whom to whom and under what circumstances. He does, as you see, refer us to other sources to find out about this marvelous group test of intelligence which he's given. He gives two sources, and in one source there are no group intelligence tests at all. In the other source there are no fewer than seven group intelligence tests which he had presented. We can't figure out which the same age. And not because they have the same genes. Unless one is assured that the test which one is using in these studies does not give different scores for different ages, or for different sexes, the studies simply hold no water. I have looked into Miss Gaw's performance test, standardized in 1925. She informs us that there are "striking" differences between the sexes in the performance test that Prof. Burt used.

Recall that Burt's is the only study that gives us data on the social and economic status of the homes in which separated twins were placed. By 1958 Burt had reported on 42 pairs of separated twins. In 1958 we were told that at least four children of high professional families —"class 1" in his study — had been reared in orphanages. But in 1966, reporting on 53 pairs, zero children of class 1 professionals had been reared in orphanages and only two of clerical workers, the next class down, had been reared in orphanages. How did four out of 42 become zero out of 53? Flat contradiction. Furthermore, Prof. Burt after 1966 mailed this marvelous data on social class to Prof. Jensen. Due to the kindness of Prof. Jensen I have a copy of the data which Prof. Burt mailed. Those data in turn contradict 1966 data, which in turn contradict the 1958 data.

If you look for mathematical impossibilities in Burt's work they assail you from every side. I really have to conclude, and I think I have presented enough evidence so that most reasonable people can follow me, that one simply cannot take Prof. Burt's numbers seriously.

I will say one thing for Prof. Burt: He was much more conscious and aware

at three months. "Brought up within a few hundred yards of one another. Attracted to each other at the age of two, but meetings were not encouraged. Told that they were twins, after the girls discovered it for themselves, gravitated to one another at school They play together quite a lot at school and during the evenings. Jessica often goes to tea with Winifred. They were never apart and wanted to sit at the same desk." This is a pair of separated twins, eight-years-old when, after this terrific separation, Shields discovered that they resembled one another.

Burton and Christopher were separated at birth. "The paternal aunts decided to take one twin each and brought them up amiably living next door to one another in the same midlands colliery village. They were constantly in and out of each others' houses." That gives you a not unfair picture of what's meant by separated twins. Separation doesn't seem terribly radical in these cases: I'm sure it's not what you had visualized.

One can also see in Shields's appendix that sometimes psychologists, for understandable reasons, have to make decisions which may affect their data. One of the tests that Shields used was the Dominoes test. The average score on this test is about 27. So he tells us about one pair of twins: Valerie got a score of 27, but her twin, Joyce, got a score of only one which was a fantastically low score. So Shields writes, "Joyce evidently did not understand the instructions; she did not give an impression of low intelligence." So the data for this pair were discarded.

Here we have a case where one twin scores 27, the other scores one, and Shields decides that the twin who scored

To page 5 of SOUTH TODAY (following Annual Report)

The Conclusions Which Have Been Drawn By Psychologists From These Data Do Not Stand

From page 4

one didn't understand the instructions. But consider what happens in another twin case, Olwen and Gwladys. Olwen had a score of four, and Gwladys had a score of two. As it happens, Olwen and Gwladys were Welsh. English was not their native language. But Shields writes, "Both twins are naturally Welsh speaking, but this does not explain their low score on the Dominos, which is in keeping with the impression on interview of low intelligence." So if you have two scores which are both abominably low, even when you know you are speaking to twins not in their native language, you conclude that they are both stupid and you keep their data. But if one scores high, and the other scores abominably low, then you use your impression to say oh, gee, one didn't understand the instructions.

If you look at the 40 pairs of separated twins who had been given the Dominoes test, for 35 pairs it is the case that Shields himself tested each member. But for the other five pairs, as it happens, two different psychologists, usually one of them being Shields, tested the two members of the twin set. Suppose we now ask, what was the average difference in score for those twins, both of whom were tested by Shields, compared with the average difference in score for those twins, each of whom was tested by a different psychologist? For twins both tested by Shields, the average difference is only 4.9. But when different psychologists tested the twins, the average difference is 13.2, and if one analyzes that difference, it turns out to be statistically significant.

What that tells us is that the degree to which identical twins resemble one another in the intelligence test depends upon whether the same psychologist is testing both twins or not. We know that intelligence tests are not precise instruments which give mechanical scores. A psychologist would be less than human if he is studying a pair of twins, and believes that twins ought to resemble one another, if he didn't exhibit this kind of biasing effect. But in all of our studies of identical twins the elementary precaution of the test being given by different testers to the two members of the twin pair has markedly? Do people always have trouble telling you apart? Do you regard yourselves as being very much alike? So they only included in their study twins who said they were alike.

They used the Stanford-Binet Intelligence Test. They had in their study a group of 100 individuals who were *non-separated* identical twins, brought up in the normal family fashion. Their monograph indicates that the correlation between age and IQ for non-separated identical twins is−.49. What that says is, the older the child is, the stupider he is! Now that's obviously nonsense. What it means is that the test had not been properly standardized so as to give a mean IQ of 100 at all ages.

If you now look at the separated identical twins, as it turns out there are 19 pairs of separated twins and seven were males, so those seven male pairs give us 14 individuals. And we can ask, what's the correlation for those 14 males between IQ and age? The answer is −.78, an enormously significant correlation. The older you are, the stupider you are.

Now it looks as if it's different for women. For if one takes the remaining 12 pairs of separated twins that were females, the correlation between age and IQ was only−.28. But there's a joker hidden there. As in all of these studies, there was a surplus of elderly women. Whenever you try to get volunteer twins, you get more women than men, and in turn you get a disproportion of elderly women. Therefore, if one discards the older women and looks just at the eight separated females within the same age range as the 14 separated males, for them the correlation between age and IQ was −.85. Not a bad match to the −.78 for the males. The point of this is the apparent resemblance between the separated twins in IQ may be due to nothing more than the fact that the twins are of the same age. Indeed, by some fairly complicated statistical procedures, this point can be made quite elegantly. Notice that you can predict a person's IQ in this study better by knowing his age than by knowing the IQ of his identical twin!

You have to look at the apparent resemblance of twins in IQ, and take out than Danish men; it means that the test is not standardized.

What about age? In Denmark, it turns out if you look at her 18 female separated twins, the correlation between age and IQ on the Wechsler is +.60, which is statistically significant. Whereas for her six males it is −.82, which is also statistically significant, but in the other direction. So if one looks at the correlation coefficients reported by Juel-Nielsen, without taking age into account, it looks as if among the nine female pairs of twins the IQ correlation is .59. But when one takes out of that statistically, as one must, the effect of age, there is essentially no IQ correlation at all between the separated twins. This is with a different test than that used by Newman, Freeman, and Holzinger, but the conclusions are very much the same.

IQ Heredity Conclusions Unsupported

There are some obvious problems with the standardization of the IQ tests used in the twin studies, which gravely affect their interpretation. The conclusions which many psychologists have drawn from these data do not, in my opinion, stand up.

It would be interesting to make the same kind of analysis of Prof. Burt's data, but that data, of course, has never been made available. We know nothing at all about the sex and ages of his twins. And we know that the data he reported about their socio-economic class was inconsistent and unreliable.

I am sorry that I have not been able to speak this morning about other kinds of IQ data used by some psychologists to support the idea that IQ tests scores are inheritable. I believe that these other kinds of data stand up even less well than the twins data, and I hope in the near future to have a detailed written analysis available.

Answers to Questions From Audience

A: If you mean by "intelligence" whatever intelligence tests happen to measure,

sions Which Have Been Drawn By Some ts From These Data Do Not Stand Up ---

markedly? Do people always have trouble telling you apart? Do you regard yourselves as being very much alike? So they only included in their study twins who said they were alike.

They used the Stanford-Binet Intelligence Test. They had in their study a group of 100 individuals who were *non-separated* identical twins, brought up in the normal family fashion. Their monograph indicates that the correlation between age and IQ for non-separated identical twins is–.49. What that says is, the older the child is, the stupider he is! Now that's obviously nonsense. What it means is that the test had not been properly standardized so as to give a mean IQ of 100 at all ages.

If you now look at the separated identical twins, as it turns out there are 19 pairs of separated twins and seven were males, so those seven male pairs give us 14 individuals. And we can ask, what's the correlation for those 14 males between IQ and age? The answer is −.78, an enormously significant correlation. The older you are, the stupider you are.

Now it looks as if it's different for women. For if one takes the remaining 12 pairs of separated twins that were females, the correlation between age and IQ was only–.28. But there's a joker hidden there. As in all of these studies, there was a surplus of elderly women. Whenever you try to get volunteer twins, you get more women than men, and in turn you get a disproportion of elderly women. Therefore, if one discards the older women and looks just at the eight separated females within the same age range as the 14 separated males, for them the correlation between age and IQ was −.85. Not a bad match to the −.78 for the males. The point of this is the apparent resemblance between the separated twins in IQ may be due to nothing more than the fact that the twins are of the same age. Indeed, by some fairly complicated statistical procedures, this point can be made quite elegantly. Notice that you can predict a person's IQ in this study better by knowing his age than by knowing the IQ of his identical twin!

You have to look at the apparent [illegible] of twins in IQ, and take out

than Danish men; it means that the test is not standardized.

What about age? In Denmark, it turns out if you look at her 18 female separated twins, the correlation between age and IQ on the Wechsler is +.60, which is statistically significant. Whereas for her six males it is −.82, which is also statistically significant, but in the other direction. So if one looks at the correlation coefficients reported by Juel-Nielsen, without taking age into account, it looks as if among the nine female pairs of twins the IQ correlation is .59. But when one takes out of that statistically, as one must, the effect of age, there is essentially no IQ correlation at all between the separated twins. This is with a different test than that used by Newman, Freeman, and Holzinger, but the conclusions are very much the same.

IQ Heredity Conclusions Unsupported

There are some obvious problems with the standardization of the IQ tests used in the twin studies, which gravely affect their interpretation. The conclusions which many psychologists have drawn from these data do not, in my opinion, stand up.

It would be interesting to make the same kind of analysis of Prof. Burt's data, but that data, of course, has never been made available. We know nothing at all about the sex and ages of his twins. And we know that the data he reported about their socio-economic class was inconsistent and unreliable.

I am sorry that I have not been able to speak this morning about other kinds of IQ data used by some psychologists to support the idea that IQ tests scores are inheritable. I believe that these other kinds of data stand up even less well than the twins data, and I hope in the near future to have a detailed written analysis available.

Answers to Questions From Audience

A: If you mean by "intelligence" whatever intelligence tests happen to measure,

kids across the country. When they looked at the results that they got with this experimental pilot version of the new test, they discovered that there was a small but statistically significant sex difference. In their sample, either the males did a little better than the females, or vice-versa. That disturbed them because it didn't make sense to them that there should be a sex difference in intelligence. So they told us what they did. They went back and looked at the individual items in the Stanford-Binet. That kind of test for children is built on the following assumption. If you give a test item to five-year-old kids, and then to six-year-old kids, and seven-year-olds, the proportion of seven-year-olds passing it ought to be greater than the proportion of six-year-olds. And the proportion of six-year-olds ought to be greater than that of five-year-olds. So, what they did was to look at the individual items and see whether they were equally age-graded for each sex. They found some individual items of which it was true that, within males, the older the kid the more likely he was to give the right answer. That wasn't true with females. And they found some items vice-versa. So they, then, perfectly rationally, threw out of the test those items. When they threw those items out of the test, then there was no longer any sex difference in the test scores between males and females.

On the next page, they tell us about social class differences. Thus it turned out that the children of professional people had average IQ's of about 125. And the children of day laborers had average IQ's of about 90. It did not occur to them to take the individual items and look to see whether they were equally age-graded for the children of professionals and for the children of the day laborers. And it didn't occur to them for a simple, profound reason. An intelligence test which did not discriminate between social classes would not be accepted by psychologists as an intelligence test.

What we mean by an intelligence test [illegible]

ence, it turns out to be statistically significant.

What that tells us is that the degree to which identical twins resemble one another in the intelligence test depends upon whether the same psychologist is testing both twins or not. We know that intelligence tests are not precise instruments which give mechanical scores. A psychologist would be less than human if he is studying a pair of twins, and believes that twins ought to resemble one another, if he didn't exhibit this kind of biasing effect. But in all of our studies of identical twins the elementary precaution of the test being given by different testers to the two members of the twin pair has never been taken.

The third study, which is the one American study in the series, was by Newman, Freeman and Holzinger. In some ways I think it's the best of the four available studies. For one thing, we know about how they got their twins. They were in Chicago and it was the Depression; they conducted a radio and newspaper campaign asking for separated twins to volunteer in the name of science. Separated identical twins. They did not have much money with which to do this research. Sometimes they got letters from Boston, sometimes from San Francisco. They did not want to take the chance of paying the transportation expenses of a possible pair of subjects to Chicago only to discover when they arrived and were blood tested that they were not really *identical* twins. What they did, as they very candidly and honestly reported, was this: Whenever people volunteered they were sent a letter and a questionnaire asking do you really resemble one another

men. Therefore, if one discards the older women and looks just at the eight separated females within the same age range as the 14 separated males, for them the correlation between age and IQ was −.85. Not a bad match to the −.78 for the males. The point of this is the apparent resemblance between the separated twins in IQ may be due to nothing more than the fact that the twins are of the same age. Indeed, by some fairly complicated statistical procedures, this point can be made quite elegantly. Notice that you can predict a person's IQ in this study better by knowing his age than by knowing the IQ of his identical twin!

You have to look at the apparent resemblance of twins in IQ, and take out of that apparent resemblance the effect of age, because two people of the same age will have similar IQ's. That's the way these IQ tests are built.

Juel-Nielsen's is the fourth separated twin study. Prof. Jensen in his review article informs us that the Wechsler, which is the test Juel-Nielsen used, had been standardized on the Danish population to give a mean of 100 and a standard deviation of 15. The problem is when you read Juel-Nielsen, she apologizes for the fact that there has never been a Danish standardization of the Wechsler test and therefore some of her data may be questionable in terms of age and sex of her subjects.

If you look at her data, it turns out for example that Danish men are brighter than Danish women. For 18 females in her sample, the mean IQ was 102; for six males it was 113. That is a statistically significant difference. It of course doesn't mean that Danish women are brighter

we know that the data he reported about their socio-economic class was inconsistent and unreliable.

I am sorry that I have not been able to speak this morning about other kinds of IQ data used by some psychologists to support the idea that IQ tests scores are inheritable. I believe that these other kinds of data stand up even less well than the twins data, and I hope in the near future to have a detailed written analysis available.

Answers to Questions From Audience

A: If you mean by "intelligence" whatever intelligence tests happen to measure, by that definition, I am saying it does not appear to be that genetics are involved in that kind of intelligence. That does not deny the possibility that there are some central cognitive capacities and abilities which we have not yet been clever enough to measure which, if we could measure, might turn out to be largely genetically determined. They might also turn out to be not genetically determined. What I am saying is that *intelligence tests scores* do not appear to be the least bit related to genetic factors. That doesn't say necessarily that it might not be possible to some day measure people as being "bright" and others "dull" and finding out that indeed this is genetically determined. The present IQ tests don't do that. I really think that distinction is important.

A: Let me cite examples from two consecutive pages from the book by Terman and Merrill in 1937, which presented the standardization of the Stanford-Binet test. What they had done was to give on an experimental basis the test to a reasonably large sample of school

social class
out that the
ple had aver
the children
IQ's of about
to take the i
see whether
for the child
the children
didn't occur
found reason
did not discri
would not b
as an intellig

What we
is a test on
sity professo
will score hi
neighbors. If
males, somet
because we
what "intelli
well as prof
whites, there
test. The te
placement o
cupational s
to use the te
somebody w
successful w
a good test.

A: Savin
cause I want
to talk ab
have been u
been some i
that they ha
policies. I w
plications on
day.

men. Therefore, if one discards the older women and looks just at the eight separated females within the same age range as the 14 separated males, for them the correlation between age and IQ was −.85. Not a bad match to the −.78 for the males. The point of this is the apparent resemblance between the separated twins in IQ may be due to nothing more than the fact that the twins are of the same age. Indeed, by some fairly complicated statistical procedures, this point can be made quite elegantly. Notice that you can predict a person's IQ in this study better by knowing his age than by knowing the IQ of his identical twin!

You have to look at the apparent resemblance of twins in IQ, and take out of that apparent resemblance the effect of age, because two people of the same age will have similar IQ's. That's the way these IQ tests are built.

Juel-Nielsen's is the fourth separated twin study. Prof. Jensen in his review article informs us that the Wechsler, which is the test Juel-Nielsen used, had been standardized on the Danish population to give a mean of 100 and a standard deviation of 15. The problem is when you read Juel-Nielsen, she apologizes for the fact that there has never been a Danish standardization of the Wechsler test and therefore some of her data may be questionable in terms of age and sex of her subjects.

If you look at her data, it turns out for example that Danish men are brighter than Danish women. For 18 females in her sample, the mean IQ was 102; for six males it was 113. That is a statistically significant difference. It of course doesn't mean that Danish women are brighter we know that the data he reported about their socio-economic class was inconsistent and unreliable.

I am sorry that I have not been able to speak this morning about other kinds of IQ data used by some psychologists to support the idea that IQ tests scores are inheritable. I believe that these other kinds of data stand up even less well than the twins data, and I hope in the near future to have a detailed written analysis available.

Answers to Questions From Audience

A: If you mean by "intelligence" whatever intelligence tests happen to measure, by that definition, I am saying it does not appear to be that genetics are involved in that kind of intelligence. That does not deny the possibility that there are some central cognitive capacities and abilities which we have not yet been clever enough to measure which, if we could measure, might turn out to be largely genetically determined. They might also turn out to be not genetically determined. What I am saying is that *intelligence tests scores* do not appear to be the least bit related to genetic factors. That doesn't say necessarily that it might not be possible to some day measure people as being "bright" and others "dull" and finding out that indeed this is genetically determined. The present IQ tests don't do that. I really think that distinction is important.

A: Let me cite examples from two consecutive pages from the book by Terman and Merrill in 1937, which presented the standardization of the Stanford-Binet test. What they had done was to give on an experimental basis the test to a reasonably large sample of school

On the next page, they tell us about social class differences. Thus it turned out that the children of professional people had average IQ's of about 125. And the children of day laborers had average IQ's of about 90. It did not occur to them to take the individual items and look to see whether they were equally age-graded for the children of professionals and for the children of the day laborers. And it didn't occur to them for a simple, profound reason. An intelligence test which did not discriminate between social classes would not be accepted by psychologists as an intelligence test.

What we mean by an intelligence test is a test on which the children of university professors, doctors, lawyers, etc., will score higher than children of their neighbors. If males score higher than females, something is wrong with our tests because we have pre-conceived ideas of what "intelligence" is. If workers score as well as professors, or blacks as well as whites, there's something wrong with the test. The tests are used to predict the placement of people in a particular occupational structure. And if you want to use the test as a prediction of whether somebody will grow up to be a wealthy, successful white middle-class male, it's a good test. It predicts that pretty well.

A: Saving that for this afternoon, because I want to do precisely that. I want to talk about how intelligence tests have been used in the past. There have been some incredible ways, to my mind, that they have been used to affect social policies. I will point to some of the implications on how they are being used today.

How Testing Harms Children

Grouping According T
Discriminatory--No Ma

By Winifred

Following is the text of a presentation at the symposium on Human Intelligence, Social Science and Social Policy by Ms. Green, director of the Atlanta-based Southeastern Public Education Program of the American Friends Service Committee.

Sgt. Moses Williams, a member of the police force in Tallulah, La., who is a black parent long active in the struggle for civil rights, told me two weeks ago: "Giving those tests and then placing children in groups is for one purpose and one purpose only and that is to segregate. Once you put a child in that low group there is no way for him to advance himself. You know some of our children may be a little slow, but they all can learn."

And two days ago, talking to a middle-class white mother in a suburban school district in the urban South, I was told: "You know Winifred, John is in his eleventh year of school and just this year, because he has two teachers who have confidence in him, is he beginning to get out of the box he was put in in elementary school because he was tracked." She went on to say that she had no worries about her own child because of the support he received at home and that he (her child) would be ok. But what, she asked, about the ones who don't get the support at home?

'Some Are Better . . .'

The effects of testing and the inevitable result of classification cut across race and class. It affects all our children, black, white, middle-class, rich and poor. Of course it is the minorities and the poor who suffer most from this device which says, no matter how sophisticated the language, SOME ARE BETTER THAN OTHERS.

I would like to spend a few minutes this afternoon telling you about some of the children, the human beings, with whom we have worked and how the current definition of intelligence and the use of testing have affected their lives and then to suggest some of the ways I think private groups (civic clubs, foundations, churches and others) can be involved in solving the problem.

We are a nation that gives lip-service to how much we value our children. How often have you heard that children are our nation's greatest resource? If we examine the facts, the statistics, and the children that we know who together make up these statistics, it is not true. We are

George L. Walker III

Children of All Races, Classes Penalized

well qualified to talk about what I have seen as the results in many Southern school systems and to tell you what the students themselves, the so-called "test users," say about them. Conversations with students in special education classes in four Georgia school districts that range

Q. Why were you assigned to Special Education classes?

A. They gave all of us a test and the dumb students was put in these kind of classes.

I ain't dumb! They gave me a test but that don't mean nothing. I just mark them

sting
hildren

Grouping According To Tests Scores Is Discriminatory--No Matter How It's Done

By Winifred Green

George L. Walker III

Children of All Races, Classes Penalized

well qualified to talk about what I have seen as the results in many Southern school systems and to tell you what the students themselves, the so-called "test users," say about them. Conversations with students in special education classes in four Georgia school districts that range

Q. Why were you assigned to Special Education classes?

A. They gave all of us a test and the dumb students was put in these kind of classes.

I ain't dumb! They gave me a test but that don't mean nothing. I just mark them

than the days of separate but unequal schools. What we have now for many black, Chicano and poor children is separate and very unequal classrooms with little if any hope of escape.

Grouping In Aiken

AFSC provided the resources for litigation against ability grouping in Aiken, S. C., in a suit now pending in federal district court. Aiken illustrates many of the absurdities of grouping, as well as its design to keep black and poor children in their place.

Grouping in Aiken is done by standardized test grades and teacher recommendations. Two teachers reported to attorneys that they had been informed by their department heads at different secondary schools to give their low groups minimum scores on the teacher recommendation form, and to give high-group students maximum scores. This could guarantee that a student would be locked into his group, whatever its level.

Another insidious aspect of grouping in Aiken is the altering of grades on students' permanent records. Students in high sections have their averages raised a letter grade on their permanent record; students in average sections retain the grades they make, and low group students are lowered a grade. Thus no one is motivated: The high groups can't lose, and the low groups can't win.

In case students in Aiken can't recognize themselves as dummies, or geniuses, it's marked for them on their report cards. High group students' cards are clearly marked "plus," and low groups are marked "minus." Thus the terms plus and minus become common references to children in those groups among both teachers and students.

Elementary students are not allowed the privilege of being smart in language arts if they are dumb in arithmetic. They are grouped at the same level in all subjects, assuring that their real skills go unchallenged at a critical period in their education.

During the first year of desegregation, the lower groups at one junior high school were segregated by sex.

Finally, there is no policy for moving students upward, and little evidence that in fact there ever is any upward movement.

Grouping Is Discriminatory

I have never seen grouping operate

OTHERS.

I would like to spend a few minutes this afternoon telling you about some of the children, the human beings, with whom we have worked and how the current definition of intelligence and the use of testing have affected their lives and then to suggest some of the ways I think private groups (civic clubs, foundations, churches and others) can be involved in solving the problem.

We are a nation that gives lip-service to how much we value our children. How often have you heard that children are our nation's greatest resource? If we examine the facts, the statistics, and the children that we know who together make up these statistics, it is not true. We are not a child-oriented nation. Children are the unrecognized, neglected and mistreated minority in our country. We need to start the process of becoming aware of how our children are mishandled.

The long struggle for equal educational opportunity has led those concerned with change in this country, and in our region particularly, to start to examine what we mean when we speak of quality education. A staff member of the AFSC once defined quality education as the process that "gives the student the basic skills which he needs for the great range of opportunities before him. . . . It (a quality education) equips him with the tools so that he . . ., not our society, can determine what he wants to do with his life and will be well prepared to take the next step toward that goal. . . ."

What Students Say

Are testing and classification procedures consistent with the aim of increasing these options for children? While I am not qualified to speak of the validity, or the fairness of testing, I am, I believe, well qualified to talk about what I have seen as the results in many Southern school systems and to tell you what the students themselves, the so-called "test users," say about them. Conversations with students in special education classes in four Georgia school districts that range from 40 to 65 percent black indicated over and over that they, the students, felt that they would function better in heterogenous classes—or as they put it, with other students not on their "level."

Here are a few of their comments:

Q. Why were you placed in Special Education?

A. I took a test and she (my teacher) said I did real bad and was assigned to this class.

Q. Do you think you have improved since being in this class?

A. No! They give us some special ways to do our work but they only confuse me.

Q. Have you told your teacher, or counselor how you feel?

A. What good would that do? They don't care. All they gonna do is start talking that stuff about a test. . . . Don't nobody care about us.

And here are comments of two other students who seemed to feel even more powerless, one of whom even viewed herself as inferior as a result of the tests:

Q. Why were you assigned to Special Education classes?

A. They gave all of us a test and the dumb students was put in these kind of classes.

I ain't dumb! They gave me a test but that don't mean nothing. I just mark them little dots as soon as I can.

The third group of students seems to reject its status in these classes. A senior in rural Georgia says:

"These classes are no good and a trick. They just trying to make us look dumb. Now when I go to class (math) we don't do that much of nothing. The teacher seems to think we are no good. She gives us a lot of what I call busy work. In order to force her to work with us we, some of us, bother her by going to her desk for help. At the end of the period we play games. Most of these games are silly; so we do a lot of complaining. She's always telling us we have a discipline problem."

I cannot be in support of testing and grouping for students such as these because our experience leads us to believe that grouping reinforces the effects of years of discriminatory treatment in the education of black children—locking them into classroom situations where curriculum, materials, teacher expectation and the resulting stigmas and hopelessness are the same as, or some predict worse

well qualified to talk about what I have seen as the results in many Southern school systems and to tell you what the students themselves, the so-called "test users," say about them. Conversations with students in special education classes in four Georgia school districts that range from 40 to 65 percent black indicated over and over that they, the students, felt that they would function better in heterogenous classes—or as they put it, with other students not on their "level."

Here are a few of their comments:

Q. Why were you placed in Special Education?

A. I took a test and she (my teacher) said I did real bad and was assigned to this class.

Q. Do you think you have improved since being in this class?

A. No! They give us some special ways to do our work but they only confuse me.

Q. Have you told your teacher, or counselor how you feel?

A. What good would that do? They don't care. All they gonna do is start talking that stuff about a test. . . . Don't nobody care about us.

And here are comments of two other students who seemed to feel even more powerless, one of whom even viewed herself as inferior as a result of the tests:

Q. Why were you assigned to Special Education classes?

A. They gave all of us a test and the dumb students was put in these kind of classes.

I ain't dumb! They gave me a test but that don't mean nothing. I just mark them little dots as soon as I can.

The third group of students seems to reject its status in these classes. A senior in rural Georgia says:

"These classes are no good and a trick. They just trying to make us look dumb. Now when I go to class (math) we don't do that much of nothing. The teacher seems to think we are no good. She gives us a lot of what I call busy work. In order to force her to work with us we, some of us, bother her by going to her desk for help. At the end of the period we play games. Most of these games are silly; so we do a lot of complaining. She's always telling us we have a discipline problem."

I cannot be in support of testing and grouping for students such as these because our experience leads us to believe that grouping reinforces the effects of years of discriminatory treatment in the education of black children—locking them into classroom situations where curriculum, materials, teacher expectation and the resulting stigmas and hopelessness are the same as, or some predict worse

Elementary students are not allowed the privilege of being smart in language arts if they are dumb in arithmetic. They are grouped at the same level in all subjects, assuring that their real skills go unchallenged at a critical period in their education.

During the first year of desegregation, the lower groups at one junior high school were segregated by sex.

Finally, there is no policy for moving students upward, and little evidence that in fact there ever is any upward movement.

Grouping Is Discriminatory

I have never seen grouping operate without students feeling that it was discriminatory. In a report we did with other organizations in 1970 ("The Status of School Desegregation, 1970"), in which we monitored 467 school districts in the South, 35 percent of the high-school and 60 percent of the elementary-school classroom segregation we found was defended on the basis of tests, usually administered for the first time with desegregation. Higher tracks were predominantly white and had white teachers; lower tracks were predominantly black and had black teachers.

We as parents, teachers and citizens may call it "special education," but the kids call it "the dumb class," or "the slow ones," or "those idiots in Mrs. Smith's class who can't read."

Grouping means that integration never really takes place. Students do not get to know one another, and in the context in which they attend school together, old stereotypes are reinforced. I might add here that when I speak of integration, I don't mean just black and white, but socio-economic integration as well.

IQ Tests As Instruments of Oppr
From Immigration Quotas to W

A Talk by DR. LEON KAMIN

On March 23, the Southern Regional Council sponsored a one-day symposium on the history of intelligence testing and its public policy implications. The principal presentation was made by Dr. Leon Kamin, chairman of the Department of Psychology at Princeton University.

During the morning session, Dr. Kamin presented findings of his extensive research into the original studies on which some American social scientists have based writings on the inheritability of intelligence.

Dr. Kamin concluded that there is no valid evidence at all to support the heredity assumption. A transcript of this portion of his presentation was published in the May-June issue of SOUTH TODAY.

The following transcript is Dr. Kamin's afternoon presentation on the history of the use of testing in the United States to support repressive public policy.

The pioneers of the intelligence testing movement in the United States were three men who, around 1910, imported from France and translated the first so-called intelligence test.

This test had been constructed by the French psychologist Alfred Binet, interestingly enough, for the purpose of detecting learning disabilities among school children, and not for the purpose of measuring a person's fixed intelligence quotient.

The three American pioneers were Lewis Terman at Stanford, Henry Goddard at Princeton, and Robert Yerkes at Harvard.

These men had some pretty well defined ideas about society and about race. Whether or not these ideas influenced their interpretations of intelligence test scores I will leave for you to decide. I'd like to read a few quotations from their writings.

Terman, in his book published in 1916 which presented the famous Stanford-Binet test, described the poor test scores of a couple of Portuguese children and then wrote the following: "The dullness seems to be racial, or at least inherent in the family stocks from which they come. The fact that one meets this type with such extraordinary frequency among Indians, Mexicans and Negroes suggests quite forcibly the whole question of racial

for the majority of cases of chronic and semi-chronic pauperism. The feeble-minded continue to multiply. Organized charities often contribute to the survival of individuals who would otherwise not be able to live and reproduce. When charity organizations help the feeble-minded float along in the social and industrial world and to produce and rear children after their kind, a doubtful service is rendered. A little psychological research would aid the united charities of any city to direct their expenditures into more profitable channels. If we would preserve our state for a class of people worthy to possess it, we must prevent as far as possible the propagation of mental degenerates. We must curtail the increasing spawn of degeneracy." You will detect perhaps a certain social and political ideology there.

Henry Goddard was invited to Princeton University in 1920 to give lectures. He had devised his own version of the Binet scale, and this is what he told the Princeton audience in 1920, after first explaining that this marvelous new test measured genetically determined intelligence. "We may perhaps be permitted to apply the principle to another problem that looms up rather large at the present time, namely socialism. Especially its extreme form of Bolshevism. Most of the arguments used by the more intelligent members of these groups are fallacious, because they ignore the mental levels [meaning IQ]. These men in their ultra humane attitudes, their desire to be fair to the workman, maintain that the great inequalities of social life are wrong and unjust.

"For example, here is a man who says, I am wearing $12 shoes, there is a laborer who is wearing $3 shoes. Why should I spend $12 while he could only afford $3? I live in a home that is artistically decorated—carpets, high priced furniture, expensive pictures, etc., and the laborer lives in a hovel. As we have said the argument is fallacious. The fact is, that workman may have a 10 year intelligence while you have a 20. To demand for him such a home as you enjoy is as absurd as it would be to insist that every laborer receive a graduate fellowship. How can there be such a thing as social equality, with this wide range of mental

ogy and Prof. Terman's.

Prof. Terman should not have been so pessimistic as he was about the possibility of convincing society that those with low IQ scores should not reproduce. Over about a 10-year period about 35 states passed compulsory sterilization laws, applied only to people with the misfortune to end up in state-supported institutions.

Robert Yerkes at Harvard entered into a small argument with Terman. Terman had pointed out that the virtue of the test was that it provided a hard scientific number which enabled one to decide who was and who wasn't feeble-minded. Well, Yerks said no, the test score certainly should be respected but, "never should a diagnosis of feeble-mindedness be made on the IQ alone. We must inquire further into the subject's economic history. What is his occupation and his pay? We must learn what we can about his immediate family. What is the economic status and occupation of the parents? When all this information has been collected, the psychologist may be of great value in getting the subject into the most suitable place in society." If his IQ was low enough, that suitable place was a public institution where he might be sterilized.

History of Immigration Laws

I think the first immediate major practical effect of intelligence testing was in an area which most people don't know about. I didn't know about it until a year ago. It turns out that intelligence tests had a great deal to do with the framing of the immigration laws of the United States. And I would like to run through briefly some of the history of that period.

As early as 1912 Prof. Goddard, the Princeton coal miner speaker, was invited to Ellis Island by the United States Public Health Service and invited to apply the new science of intelligence testing to immigrants from Europe. Goddard reported in a scientific paper that he had given the intelligence test to a representative sample of European immigrants. He discovered that 83 percent of the Jews trying to enter the U. S. were feeble-minded, that 80 percent of the Hungarians were feeble-minded, that 79 percent of the Italians were feeble-minded, and that 87 percent of the Russians were

s As Instruments of Oppression--
mmigration Quotas to Welfare

A Talk by DR. LEON KAMIN

for the majority of cases of chronic and semi-chronic pauperism. The feeble-minded continue to multiply. Organized charities often contribute to the survival of individuals who would otherwise not be able to live and reproduce. When charity organizations help the feeble-minded float along in the social and industrial world and to produce and rear children after their kind, a doubtful service is rendered. A little psychological research would aid the united charities of any city to direct their expenditures into more profitable channels. If we would preserve our state for a class of people worthy to possess it, we must prevent as far as possible the propagation of mental degenerates. We must curtail the increasing spawn of degeneracy." You will detect perhaps a certain social and political ideology there.

Henry Goddard was invited to Princeton University in 1920 to give lectures. He had devised his own version of the Binet scale, and this is what he told the Princeton audience in 1920, after first explaining that this marvelous new test measured genetically determined intelligence. "We may perhaps be permitted to apply the principle to another problem that looms up rather large at the present time, namely socialism. Especially its extreme form of Bolshevism. Most of the arguments used by the more intelligent members of these groups are fallacious, because they ignore the mental levels [meaning IQ]. These men in their ultra humane attitudes, their desire to be fair to the workman, maintain that the great inequalities of social life are wrong and unjust.

"For example, here is a man who says, I am wearing $12 shoes, there is a laborer who is wearing $3 shoes. Why should I spend $12 while he could only afford $3? I live in a home that is artistically decorated—carpets, high priced furniture, expensive pictures, etc., and the laborer lives in a hovel. As we have said the argument is fallacious. The fact is, that workman may have a 10 year intelligence while you have a 20. To demand for him such a home as you enjoy is as absurd as it would be to insist that every laborer receive a graduate fellowship. How can there be such a thing as social

ogy and Prof. Terman's.

Prof. Terman should not have been so pessimistic as he was about the possibility of convincing society that those with low IQ scores should not reproduce. Over about a 10-year period about 35 states passed compulsory sterilization laws, applied only to people with the misfortune to end up in state-supported institutions.

Robert Yerkes at Harvard entered into a small argument with Terman. Terman had pointed out that the virtue of the test was that it provided a hard scientific number which enabled one to decide who was and who wasn't feeble-minded. Well, Yerks said no, the test score certainly should be respected but, "never should a diagnosis of feeble-mindedness be made on the IQ alone. We must inquire further into the subject's economic history. What is his occupation and his pay? We must learn what we can about his immediate family. What is the economic status and occupation of the parents? When all this information has been collected, the psychologist may be of great value in getting the subject into the most suitable place in society." If his IQ was low enough, that suitable place was a public institution where he might be sterilized.

History of Immigration Laws

I think the first immediate major practical effect of intelligence testing was in an area which most people don't know about. I didn't know about it until a year ago. It turns out that intelligence tests had a great deal to do with the framing of the immigration laws of the United States. And I would like to run through briefly some of the history of that period.

As early as 1912 Prof. Goddard, the Princeton coal miner speaker, was invited to Ellis Island by the United States Public Health Service and invited to apply the new science of intelligence testing to immigrants from Europe. Goddard reported in a scientific paper that he had given the intelligence test to a representative sample of European immigrants. He discovered that 83 percent of the Jews trying to enter the U. S. were feeble-minded, that 80 percent of the Hungarians were feeble-minded, that 79 percent of the Italians were feeble-minded, and

cluded, but the total number of immigrants was still wide open.

Then one sees the first straw in the wind. Terman, Yerkes and Goddard and many of the leading biologists and psychologists were members of an organization called the Eugenics Research Association. They were concerned with improving human breeding stock by cutting off the defective germ plasm of the feeble-minded, etc. The editor of their Journal, *The Eugenical News*, was a biologist named Dr. Harry Laughlin. As early as 1917 he wrote the following under the heading "The New Immigration Law": "Recently the science of psychology has developed to a high stage of precision that branch of its general subject devoted to the testing of individuals for natural excellence in mental and temperamental qualities. When the knowledge of the existence of this science becomes generally known in Congress, that body will then be expected to apply the direct and logical test for the qualities which we seek to measure in immigrants."

World War I and IQ Testing

The small cloud on the horizon is no larger than a man's fist, but it's going to grow and grow. Later in 1917, the U. S. entered World War I and for the first time in history, literally millions of individuals were subjected to intelligence testing. The head of the intelligence testing pogram in the Army was now Col. Robert M. Yerkes.

In 1918 a small distinguished group of psychologists and biologists from the eastern seaboard founded something called the Galton Society which met in the American Museum of Natural History in New York City once a month. It consisted at no one time of more than 25 distinguished scientists. They provided, as a free service, to the government and to private organizations all kinds of helpful hints about eugenical practices.

In 1920 there was an enormous influx of experimental psychologists into the Eugenics Research Association. All at the same time in a given year, a tremendous number of distinguished experimental psychologists and biologists

fined ideas about society and about race. Whether or not these ideas influenced their interpretations of intelligence test scores I will leave for you to decide. I'd like to read a few quotations from their writings.

Terman, in his book published in 1916 which presented the famous Stanford-Binet test, described the poor test scores of a couple of Portuguese children and then wrote the following: "The dullness seems to be racial, or at least inherent in the family stocks from which they come. The fact that one meets this type with such extraordinary frequency among Indians, Mexicans and Negroes suggests quite forcibly the whole question of racial differences in mental traits will have to be taken up anew and by experimental methods."

The writer predicts that when this is done "there will be discovered enormously significant racial differences in general intelligence, differences which cannot be wiped out by any scheme of mental culture. Children of this group should be segregated in special classes. They cannot master abstractions, but they can often be made efficient workers. There is no possibility at present of convincing society that they should not be allowed to reproduce. They constitute a grave menace because of their unusually prolific breeding." This, from the founder of the American intelligence testing movement.

"The Menace of Feeble-Mindedness"

The next year writing under the heading "The Menace of Feeble-Mindness," Prof. Terman wrote, "Only recently have we begun to recognize how serious a menace it is to the social, economic and moral welfare of the state. It is responsible

inequalities of social life are wrong and unjust.

"For example, here is a man who says, I am wearing $12 shoes, there is a laborer who is wearing $3 shoes. Why should I spend $12 while he could only afford $3? I live in a home that is artistically decorated—carpets, high priced furniture, expensive pictures, etc., and the laborer lives in a hovel. As we have said the argument is fallacious. The fact is, that workman may have a 10 year intelligence while you have a 20. To demand for him such a home as you enjoy is as absurd as it would be to insist that every laborer receive a graduate fellowship. How can there be such a thing as social equality, with this wide range of mental capacity?

"As for an equal distribution of the wealth of the world, that is equally absurd. The man with intelligence has spent his money wisely and saved until he has enough to provide for his needs in case of sickness, while a man of low intelligence, no matter how much money he would have earned, would have spent much of it foolishly and would never have anything ahead. It is said that during the past year the coal miners in certain parts of the country have earned more money than the coal operators. Yet today when the mines shut down for a time, those people are the first to suffer. They did not save anything, although their whole life has taught them that mining is an irregular thing, and that when they were having plenty of work, they should save against the days when they do not have work."

But low IQ coal miners, according to Prof. Goddard, don't have enough intelligence to salt away all that extra money they're making, etc. You may detect a similarity between Prof. Goddard's ideol-

of the immigration laws of the United States. And I would like to run through briefly some of the history of that period.

As early as 1912 Prof. Goddard, the Princeton coal miner speaker, was invited to Ellis Island by the United States Public Health Service and invited to apply the new science of intelligence testing to immigrants from Europe. Goddard reported in a scientific paper that he had given the intelligence test to a representative sample of European immigrants. He discovered that 83 percent of the Jews trying to enter the U. S. were feeble-minded, that 80 percent of the Hungarians were feeble-minded, that 79 percent of the Italians were feeble-minded, and that 87 percent of the Russians were feeble-minded. Science and psychology score again. And he wrote in 1917 "The number of aliens deported because of feeble-mindedness increased approximately 350 percent in 1913 and 57 percent in 1914 over what it had been. This was due chiefly to the efforts of the physicians who were inspired by the belief that mental tests could be used for the detection of feeble-minded aliens."

Until 1875 there were no federal immigration laws, whatsoever. Anybody who wanted to came. In 1875, the first federal law was passed. It did not limit the number of immigrants, but it did exclude certain classes of individuals. And the first exclusion in 1875 consisted of three types of people—coolies, convicts and foreign prostitutes. In 1882, lunatics and idiots were added, and as one goes through the years more classes. So by 1903 persons who had had two or more attacks of insanity at any time previously were added. The point is that up until the outbreak of World War I, there was in principle unlimited immigration. Certain defective classes of individuals were ex-

inequalities of social life are wrong and unjust.

"For example, here is a man who says, I am wearing $12 shoes, there is a laborer who is wearing $3 shoes. Why should I spend $12 while he could only afford $3? I live in a home that is artistically decorated—carpets, high priced furniture, expensive pictures, etc., and the laborer lives in a hovel. As we have said the argument is fallacious. The fact is, that workman may have a 10 year intelligence while you have a 20. To demand for him such a home as you enjoy is as absurd as it would be to insist that every laborer receive a graduate fellowship. How can there be such a thing as social equality, with this wide range of mental capacity?

"As for an equal distribution of the wealth of the world, that is equally absurd. The man with intelligence has spent his money wisely and saved until he has enough to provide for his needs in case of sickness, while a man of low intelligence, no matter how much money he would have earned, would have spent much of it foolishly and would never have anything ahead. It is said that during the past year the coal miners in certain parts of the country have earned more money than the coal operators. Yet today when the mines shut down for a time, those people are the first to suffer. They did not save anything, although their whole life has taught them that mining is an irregular thing, and that when they were having plenty of work, they should save against the days when they do not have work."

But low IQ coal miners, according to Prof. Goddard, don't have enough intelligence to salt away all that extra money they're making, etc. You may detect a similarity between Prof. Goddard's ideology and the immigration laws of the United States. And I would like to run through briefly some of the history of that period.

As early as 1912 Prof. Goddard, the Princeton coal miner speaker, was invited to Ellis Island by the United States Public Health Service and invited to apply the new science of intelligence testing to immigrants from Europe. Goddard reported in a scientific paper that he had given the intelligence test to a representative sample of European immigrants. He discovered that 83 percent of the Jews trying to enter the U. S. were feeble-minded, that 80 percent of the Hungarians were feeble-minded, that 79 percent of the Italians were feeble-minded, and that 87 percent of the Russians were feeble-minded. Science and psychology score again. And he wrote in 1917, "The number of aliens deported because of feeble-mindedness increased approximately 350 percent in 1913 and 57 percent in 1914 over what it had been. This was due chiefly to the efforts of the physicians who were inspired by the belief that mental tests could be used for the detection of feeble-minded aliens."

Until 1875 there were no federal immigration laws, whatsoever. Anybody who wanted to came. In 1875, the first federal law was passed. It did not limit the number of immigrants, but it did exclude certain classes of individuals. And the first exclusion in 1875 consisted of three types of people—coolies, convicts and foreign prostitutes. In 1882, lunatics and idiots were added, and as one goes through the years more classes. So by 1903 persons who had had two or more attacks of insanity at any time previously were added. The point is that up until the outbreak of World War I, there was in principle unlimited immigration. Certain defective classes of individuals were excluded.

of psychologists and biologists from the eastern seaboard founded something called the Galton Society which met in the American Museum of Natural History in New York City once a month. It consisted at no one time of more than 25 distinguished scientists. They provided, as a free service, to the government and to private organizations all kinds of helpful hints about eugenical practices.

In 1920 there was an enormous influx of experimental psychologists into the Eugenics Research Association. All at the same time in a given year, a tremendous number of distinguished experimental psychologists and biologists in the country joined the Eugenics Research Association. The secretary of the association, Dr. Laughlin, got a new job. He remained secretary of the Association but he was also appointed "Expert Eugenics Agent" for the House Committee on Immigration and Naturalization of the U. S. Congress.

In 1921, the National Academy of Science published, under the editorship of Col. Yerkes, an enormous volume which contained within it the summary of all the intelligence testing data gathered on draftees in World War I.

Most people know that that volume demonstrated massively for the first time that blacks scored much lower than whites in intelligence tests. But what almost nobody seems to know was that the immediate practical application of this data was not to a black-white question but to the question of immigration. There was a special chapter in this volume edited by Yerkes, which presented the data on the intelligence test scores of draftees who were immigrants and who had been born in different European countries.

To page 8

Profound Economic Conflicts, Pr
Racism Behind Misuse of IQ T

From page 7

The U. S. Army test scores could be classified in terms of A, B, C, D, E. If you are really stupid you are E. What that shows is that the average intelligence of draftees who are immigrants from different European countries differ. It turns out that immigrants coming from England, Holland, Denmark, Scotland, Germany, Switzerland, etc. seem quite bright. The Russians, the Italians and the Poles are all just plain stupid. The Poles did not score higher than the blacks. Now in describing this set of interesting data, the writer used the value-free statement, "The Slavic and Latin countries stand low."

"The Passing of the Great Race"

Madison Grant, president of the Galton Society had published a very influential racist book called *The Passing of the Great Race*, in which he divided all of Europe into Mediterranean, Alpine and Nordic stocks. Nordics were later re-christened Aryan by other eugenic authorities. In any event, Grant had made the observation that all that was good, beautiful and true in culture and civilization had come to us from the Nordics and nothing much from the Latin and Slavic countries. Here we see proof that the intelligence of the Latins and Slavs is indeed visibly dull.

In 1922, Col. Yerkes had a civilian job in the National Research Council in Washington. The National Research Council under his leadership set up a special commission of scientists called the Commission for the Scientific Study of Problems of Human Migration. The National Research Council began to support research relevant to problems of human migration. The first research it supported was that of Prof. Carl Brigham, at that time an assistant professor at Princeton University in the psychology department. In 1923, the Princeton University Press published Brigham's work entitled, *A Study of American Intelligence*. What Brigham did was take the Army immigrant data and demonstrate that the longer an immigrant had been in the United States, the higher was his IQ.

We must assume that he assumed the test measured native intelligence. The psychologists had built it to measure native intelligence. How can it be that of Europe is comparatively free from this taint."

The idea was that the lower classes breed too much and the upper classes don't breed enough. So, since intelligence is inherited, over the years the average intelligence is going to sink and sink until eventually it gets below zero. Here in the U. S. it's going to be worse, because not only do we have stupid immigrants, but we have stupid blacks as well.

Brigham wrote, "But the decline of American intelligence will be more rapid than the decline of intelligence of European national groups, owing to the presence here of a Negro. The steps that should be taken must of course be dictated by science and not by political expedience. Immigration should not only be restricted but highly selective. And the revision of the immigration and naturalization laws will only afford a slight relief from our present difficulties. The really important steps are those looking toward the prevention of the continued propagation of the defective strains in the present population."

For those of you who missed the point, Prof. Brigham advocated the compulsory sterilization of the "defective." For those who want to know what happened to Prof. Brigham, he went on to a job as secretary to the College Entrance Examination Board. Also the Scholastic Aptitude Test was devised by Prof. Brigham. His book was not without influence.

Congressional Hearings

In 1923, the Eugenics Research Association elected as its chairman a non-scientist. That layman was the Hon. Albert Johnson, chairman of the House Committee on Immigration and Naturalization. The committee began to have hearings. I have read through the *Congressional Records* and hearing testimony in some detail. I'll just read a couple of excerpts. Dr. Arthur Sweeney testifying before the committee in 1923, said, "We have been overrun by a horde of the unfit. We've had no adequate means of determining who are the unfit. While we can measure objectively the physical qualifications of the immigrants, we have had no yardstick with which to perform an accurate estimate of the intellectual side. Psychological tests in the Army have furnished us with the necessary yardstick. Immigration from eastern and southern Europe is most undesirable. We can gauge the desirability of immigrants by the relative proportion of these in the northern European group. From this source the stream of intelligent citizenship is polluted. Who are the unfit? The groups at the lower end of the psychological scale."

Francis Kinnicutt, testifying before a Senate Committee on Immigration and Naturalization in 1923, said, "The large proportion of this immigration from southern and eastern Europe comes from Poland or from Russia. The immigration from the last two countries consists largely of the Hebrew elements. Some of their labor unions are among the most radical in the whole country. The recent Army tests show that all three of these classes ranked far below the average in intelligence." Here, he was referring to Brigham's book, *A Study of American Intelligence*. Col. Yerkes also vouches for Brigham's book and speaks in the highest terms of the author, who, by this time, is assistant professor of psychology at Princeton University.

Also testifying at these hearings was Madison Grant, who said, "The country at large has been greatly impressed by the results of the Army intelligence tests, carefully analyzed by Yerkes and Brigham. The experts who have analyzed the statistics and who have tested the tests believe that the tests give as accurate a measurement of intelligence as is possible at the present time. The questions were selected with a view to measuring innate ability. Experts have told us that had mental testing been in operation, over 6 million Americans now living in this country, free to become the fathers and mothers of future Americans, would never have been admitted."

New Restrictive Immigration Laws

In 1921, as the result of an intense national debate, temporary legislation limiting the number of immigrants was passed by Congress for the first time. According to the law, the number of immigrants allowed into the United States from any European country depended upon the number of people from that European country already residing in the U. S. This temporary law assigned to each European country an immigration quota based on three percent of its population residing in the U. S. Even though the temporary law passed in 1921, the popuiation statistics were based on the 1910 Census.

Just so that you don't miss the point, the permanent law which was passed in

d Economic Conflicts, Profound sm Behind Misuse of IQ Tests

of Europe is comparatively free from this taint."

The idea was that the lower classes breed too much and the upper classes don't breed enough. So, since intelligence is inherited, over the years the average intelligence is going to sink and sink until eventually it gets below zero. Here in the U. S. it's going to be worse, because not only do we have stupid immigrants, but we have stupid blacks as well.

Brigham wrote, "But the decline of American intelligence will be more rapid than the decline of intelligence of European national groups, owing to the presence here of a Negro. The steps that should be taken must of course be dictated by science and not by political expedience. Immigration should not only be restricted but highly selective. And the revision of the immigration and naturalization laws will only afford a slight relief from our present difficulties. The really important steps are those looking toward the prevention of the continued propagation of the defective strains in the present population."

For those of you who missed the point, Prof. Brigham advocated the compulsory sterilization of the "defective." For those who want to know what happened to Prof. Brigham, he went on to a job as secretary to the College Entrance Examination Board. Also the Scholastic Aptitude Test was devised by Prof. Brigham. His book was not without influence.

Congressional Hearings

In 1923, the Eugenics Research Association elected as its chairman a non-scientist. That layman was the Hon. Albert Johnson, chairman of the House Committee on Immigration and Naturalization. The committee began to have hearings. I have read through the *Congressional Records* and hearing testimony in some detail. I'll just read a couple of excerpts. Dr. Arthur Sweeney testifying before the committee in 1923, said, "We have been overrun by a horde of the unfit. We've had no adequate means of determining who are the unfit. While we can measure objectively the physical qualifications of the immigrants, we have had no yardstick with which to perform an accurate estimate of the intellectual side. Psychological tests in the Army have furnished us with the necessary yardstick. Immigration from eastern and southern Europe is most undesirable. We can gauge the desirability of immigrants northern European group. From this source the stream of intelligent citizenship is polluted. Who are the unfit? The groups at the lower end of the psychological scale."

Francis Kinnicutt, testifying before a Senate Committee on Immigration and Naturalization in 1923, said, "The large proportion of this immigration from southern and eastern Europe comes from Poland or from Russia. The immigration from the last two countries consists largely of the Hebrew elements. Some of their labor unions are among the most radical in the whole country. The recent Army tests show that all three of these classes ranked far below the average in intelligence." Here, he was referring to Brigham's book, *A Study of American Intelligence*. Col. Yerkes also vouches for Brigham's book and speaks in the highest terms of the author, who, by this time, is assistant professor of psychology at Princeton University.

Also testifying at these hearings was Madison Grant, who said, "The country at large has been greatly impressed by the results of the Army intelligence tests, carefully analyzed by Yerkes and Brigham. The experts who have analyzed the statistics and who have tested the tests believe that the tests give as accurate a measurement of intelligence as is possible at the present time. The questions were selected with a view to measuring innate ability. Experts have told us that had mental testing been in operation, over 6 million Americans now living in this country, free to become the fathers and mothers of future Americans, would never have been admitted."

New Restrictive Immigration Laws

In 1921, as the result of an intense national debate, temporary legislation limiting the number of immigrants was passed by Congress for the first time. According to the law, the number of immigrants allowed into the United States from any European country depended upon the number of people from that European country already residing in the U. S. This temporary law assigned to each European country an immigration quota based on three percent of its population residing in the U. S. Even though the temporary law passed in 1921, the population statistics were based on the 1910 Census.

Just so that you don't miss the point, the permanent law which was passed in the intelligence of children of southeastern European immigrants whose children had been born in this country had gone to schools in this country. If the stupidity is really genetic, then the children of Polish, Russian and Italian immigrants should be stupid as well.

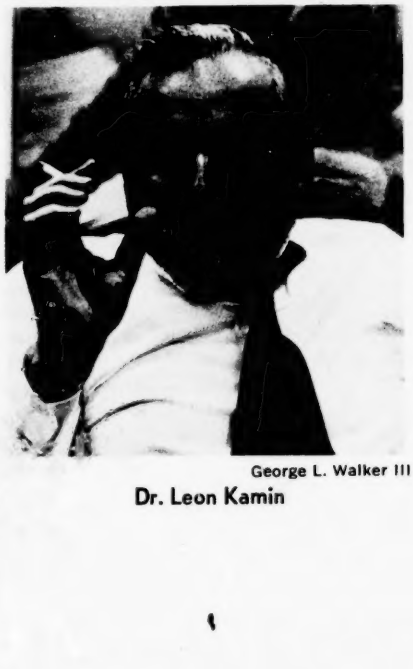
Needless to say, Prof. Hirsch discovered that indeed the children of Russian, Polish, Italian immigrants were also stupid. Let me read to you from his study. Remember this is in a professional, psychological journal. "The immigration law passed by Congress is one of most hopeful signs. That part of the law which has to do with non-quota immigrants should be modified." (You see, there was no quota for immigrants either from Canada or from Mexico.) "All mental testing upon children of Spanish Mexican descent has shown that the average intelligence of this group is even lower than the average intelligence of Portuguese and Negro children in this study. Yet Mexicans are flowing into the country. In our immigration from Canada, we are getting the less intelligent, working-class people. The increase in the number of French Canadians is alarming. Whole New England villages and towns are filled with them. The average intelligence of the French Canadian group in our data approaches the level of the average Negro intelligence. I have seen gatherings of the foreign born in which narrow and sloping foreheads were the rule. In every face there was something wrong. Lips thick, mouth coarse, chin poorly formed, sugar loaf heads, goose-bill noses, . . . Immigration officials report vast trouble in extracting the truth from certain brunette nationalities."

I think what happened is fairly summarized in the presidential address of Frank Babbott to the Eugenics Research Association in 1927. By that time the battle was over and the national origin quotas had been passed. Speaking to the biologists and psychologists in the Association, Babbott said the following: "Eugenics has made its strongest appeal to me by its influence on immigration. This is an indirect result of eugenics, but it comes as a natural development of research on the part of people like yourselves. It is possible that restriction of immigration would have come without the aid of our society, but I doubt if it would have come so soon or so permanently."

National Research Council began to support research relevant to problems of human migration. The first research it supported was that of Prof. Carl Brigham, at that time an assistant professor at Princeton University in the psychology department. In 1923, the Princeton University Press published Brigham's work entitled, *A Study of American Intelligence*. What Brigham did was take the Army immigrant data and demonstrate that the longer an immigrant had been in the United States, the higher was his IQ.

We must assume that he assumed the test measured native intelligence. The psychologists had built it to measure native intelligence. How can it be that immigrants that have been in this country a long time are so much brighter than immigrants who haven't been in this country a long time? Prof. Brigham concluded that this had nothing to do with knowing the English language or knowing American culture—no. The immigrants, during the early 1900s, had arrived from northwestern Europe, full of Nordic blood, whereas the immigrants who arrived from 1913 to 1918 were from southeastern Europe and consisted of Poles, Italians, Russians, Jews, etc., immigrants with very little Nordic blood.

From this Prof. Brigham concluded that there was a clear parallel between the proportion of Nordic blood and the intelligence of the immigrants. That left a certain policy recommendation. Let me read to you the concluding passages of Prof. Brigham's book. (The foreword in this book was written by Col. Yerkes.) Brigham says, "We must face a possibility of racial admixture here that is infinitely worse than that faced by any European country today. For we are incorporating the Negro into our racial stock, while all hearings. I have read through the *Congressional Records* and hearing testimony in some detail. I'll just read a couple of excerpts. Dr. Arthur Sweeney testifying before the committee in 1923, said, "We have been overrun by a horde of the unfit. We've had no adequate means of determining who are the unfit. While we can measure objectively the physical qualifications of the immigrants, we have had no yardstick with which to perform an accurate estimate of the intellectual side. Psychological tests in the Army have furnished us with the necessary yardstick. Immigration from eastern and southern Europe is most undesirable. We can gauge the desirability of immigrants by the relative proportion of those in the A and B classes. We strenuously object to the immigration from Italy, Russia, Poland, Greece, and Turkey. The Slavic and Latin countries show a marked contrast in intelligence with the western and



George L. Walker III

Dr. Leon Kamin

In 1921, as the result of an intense national debate, temporary legislation limiting the number of immigrants was passed by Congress for the first time. According to the law, the number of immigrants allowed into the United States from any European country depended upon the number of people from that European country already residing in the U. S. This temporary law assigned to each European country an immigration quota based on three percent of its population residing in the U. S. Even though the temporary law passed in 1921, the population statistics were based on the 1910 Census.

Just so that you don't miss the point, the permanent law which was passed in 1924 reduced the percentage of the residing European population to two percent, based not on the 1910 census, certainly not on the 1920 census, not on the 1900 census but on the *1890 census*, when all the immigrant blood was Nordic. If one reads the congressional debate, nobody fooled anybody. The use of the 1890 census was explicitly stated to be an exclusionary racist policy—to keep the inferior blood from southeastern Europe away from our fair shores, while making it possible for northwestern Europeans to immigrate. This was largely rationalized on the basis that the tests of innate ability developed by psychologists had clearly demonstrated, in study after study, that Italians, Poles, Russians and Jews scored 25 or 30 IQ points lower than the Nordics from northwestern Europe.

You can't appreciate the full flavor of the "scientific findings" without reading the literature. Prof. Hirsch conducted a study in 1926, supported by the National Research Council, and published in the professional journal, *Genetic Psychology Monographs*, that raised the question of

Frank Babbo
Association
battle was o
quotas had b
biologists an
sociation, B
"Eugenics ha
to me by it
This is an
but it come
of research
yourselves.
of immigrati
the aid of o
would have
nently."

Today'

The majo
country in th
of immigrat
economic co
The major
day, our pol
welfare me
major econo
racism. The
tific biologis
prepared to
Committee o
zation. I'm a
age of biolog
prepared to

The data
now presenti
is no more v
by their sci
1920s on the
ligence amon

That data
was absurd.
"innate inte
today's data
50 years.

hearings. I have read through the *Congressional Records* and hearing testimony in some detail. I'll just read a couple of excerpts. Dr. Arthur Sweeney testifying before the committee in 1923, said, "We have been overrun by a horde of the unfit. We've had no adequate means of determining who are the unfit. While we can measure objectively the physical qualifications of the immigrants, we have had no yardstick with which to perform an accurate estimate of the intellectual side. Psychological tests in the Army have furnished us with the necessary yardstick. Immigration from eastern and southern Europe is most undesirable. We can gauge the desirability of immigrants by the relative proportion of those in the A and B classes. We strenuously object to the immigration from Italy, Russia, Poland, Greece, and Turkey. The Slavic and Latin countries show a marked contrast in intelligence with the western and

George L. Walker III

Dr. Leon Kamin

In 1921, as the result of an intense national debate, temporary legislation limiting the number of immigrants was passed by Congress for the first time. According to the law, the number of immigrants allowed into the United States from any European country depended upon the number of people from that European country already residing in the U. S. This temporary law assigned to each European country an immigration quota based on three percent of its population residing in the U. S. Even though the temporary law passed in 1921, the population statistics were based on the 1910 Census.

Just so that you don't miss the point, the permanent law which was passed in 1924 reduced the percentage of the residing European population to two percent, based not on the 1910 census, certainly not on the 1920 census, not on the 1900 census but on the *1890 census,* when all the immigrant blood was Nordic. If one reads the congressional debate, nobody fooled anybody. The use of the 1890 census was explicitly stated to be an exclusionary racist policy—to keep the inferior blood from southeastern Europe away from our fair shores, while making it possible for northwestern Europeans to immigrate. This was largely rationalized on the basis that the tests of innate ability developed by psychologists had clearly demonstrated, in study after study, that Italians, Poles, Russians and Jews scored 25 or 30 IQ points lower than the Nordics from northwestern Europe.

You can't appreciate the full flavor of the "scientific findings" without reading the literature. Prof. Hirsch conducted a study in 1926, supported by the National Research Council, and published in the professional journal, *Genetic Psychology Monographs,* that raised the question of Frank Babbott to the Eugenics Research Association in 1927. By that time the battle was over and the national origin quotas had been passed. Speaking to the biologists and psychologists in the Association, Babbott said the following: "Eugenics has made its strongest appeal to me by its influence on immigration. This is an indirect result of eugenics, but it comes as a natural development of research on the part of people like yourselves. It is possible that restriction of immigration would have come without the aid of our society, but I doubt if it would have come so soon or so permanently."

Today's Data — "Absurd"

The major domestic issue facing the country in the early 1920s was the issue of immigration. It involved profound economic conflicts and profound racism. The major issue facing the country today, our politicians tell us, is "the great welfare mess," which again involves major economic conflict and profound racism. There was no shortage of scientific biologists and psychologists in 1920 prepared to act as teachers to the House Committee on Immigration and Naturalization. I'm afraid there will be no shortage of biologists, psychologists and others prepared to act as teachers today.

The data which such teachers are now presenting to policymakers, I think, is no more valid than the data presented by their scientific predecessors of the 1920s on the differences in innate intelligence among the racial stocks of Europe.

That data, as all would agree today, was absurd. It had nothing to do with "innate intelligence." In my opinion, today's data will look just as absurd in 50 years.